(11) Application No. AU 199748747 B2 (12) PATENT (10) Patent No. 730900 (19) AUSTRALIAN PATENT OFFICE (54) Improvements in or relating to starch content of plants International Patent Classification(s) (51)⁶ C12N 009/10 C12N 015/82 C12N 015/54 A01H 005/00 C12Q 001/68 C12N 001/21 (22) Application Date: 1997 .11 .04 199748747 Application No: (21)WIPO No: W098/20145 (87)(30)Priority Data (33) Date Country (32)Number (31)GΒ 1996 .11 .05 9623095 1998 .05 .29 (43)Publication Date: Publication Journal Date: 1998 .07 .23 (43) Accepted Journal Date: 2001 .03 .15 (44) (71) Applicant(s) National Starch and Chemical Investment Holding Corporation inventor(s) (72)Safford Richard Alan Jobling; Stephen Agent/Attorney (74) COLLISON and CO.GPO Box 2556, ADELAIDE 5001 (56)Related Art BIOLOGY. 20(5):809-819 PLANT MOLECULAR THE PLANT JOURNAL. 7(1):3-5 PLANT MOLECULAR BIOLOGY 30:97-108

11



.T)

(51) International Patent Classification 6: WO 98/20145 (11) International Publication Number: C12N 15/82, 9/10, C12Q 1/68, C12N 1/21, 15/54, A01H 5/90 A2 (43) International Publication Date: 14 May 1998 (14.05.98) (31) International Application Number: \ PCT/GB97/03032 (83) Designated States: AU, BR, CA, JP, KR, US, Buropean patent (AT, BB, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). (22) International Filing Date: 4 November 1997 (04.11.97) (30) Priority Data: 9623095.8 5 November 1996 (05,11.96) Without international search report and to be republished upon receipt of that report. (71) Applicant (for all designated States except US): NATIONAL STARCH AND CHEMICAL INVESTMENT HOLDENG CORPORATION (US/US): Suits 77, 601 Silverside Road Wilmington, DE 19809 (US). 341, 601 Silverside Road SEC 1985 O 104 104 (73) Inventors; and
(75) Inventors; and
(75) Inventors/Applicants (for US only): JOBLING, Stephen, Alan
[GB/GB]; 19 Burwell Road, Baton Socon, Hantingdon
PB19 3QQ (GB), SAFFORD, Richard [GB/GB]; 10 Pursess
Close, Bedford, Bedfordshire MK41 SRN (GB). EM OF (74) Agent: KEITH W NASH & CO.; 90-92 Regent Street, Cambridge CB2 IDP (GB). (54) THIS: IMPROVEMENTS IN OR RELATING TO STARCH CONTENT OF PLANTS 1kb 赤缸 TRACE 27 TRACE. 11 (8 7 ... is the ... is the (1911.2mg MALIFIEL SELECT (57) Abstract Disclosed is a nucleic acid sequence encoding a polypeptide having starts branching enzyme (SBE) activity, the encoded polypeptide comprising an effective portion of the smino acid sequence shown in Figure 4 or Figure 13.

Title: Improvements in or Relating to Starch Content of Plants

Field of the Invention

This invention relates to novel nucleic acid sequences, vectors and host cells comprising the nucleic acid sequence(s), to polypeptides encoded thereby, and to a method of altering a host cell by introducing the nucleic acid sequence(s) of the invention.

Background to the Invention

Starch consists of two main polysaccharides, amylose and amylopectin. Amylose is a linear polymer containing α -1,4 linked glucose units, while amylopectin is a highly branched polymer consisting of a α -1,4 linked glucon backbone with α -1,6 linked glucon branches. In most plant storage reserves amylopectin consistutes about 75% of the starch content. Amylopectin is synthesized by the concerted action of soluble starch synthase and starch branching enzyme [α -1,4 glucan: α -1,4 glucan 6-glycosyltransferase, EC 2.4.1.18]. Starch branching enzyme (SBE) hydrolyses α -1.4 linkages and rejoins the cleaved glucan, via an α -1.6 linkage, to an acceptor chain to produce a branched structure. The physical properties of starch are strongly affected by the relative abundance of amylose and amylopectin, and SBE is therefore a crucial enzyme in determining both the quantity and quality of starches produced in plant systems.

Starches are commercially available from several plant sources including maize, potato and cassava. Each of these starches has unique physical characteristics and properties and a variety of possible industrial uses. In maize there are a number of naturally occurring mutants which have altered starch composition such as high amylopectin types ("waxy" starches) or high amylose starches but in potato and cassava no such mutants exist on a commercial basis as yet.

Genetic modification offers the possibility of obtaining new starches which may have novel and potentially useful characteristics. Most of the work to date has involved potato plants because they are amenable to genetic manipulation i.e. they can be transformed using Agrobacterium and regenerated easily from tissue culture. In addition many of the genes involved in starch biosynthesis have been cloned from potato and thus are available as targets for genetic manipulation, for example, by antisense inhibition of expression or sense suppression.

Cassava (Manihot esculenta L. Crantz) is an important crop in the tropics, where its starch-filled roots are used both as a food source and increasingly as a source of starch. Cassava is a high yielding perennial crop that can grow on poor soils and is also tolerant of drought. Cassava starch being a root-derived starch has properties similar but not identical to potato starch and is composed of 20-25% amylose and 75-80% amylopectin (Rickard et al., 1991. Trop. Sci. 31, 189-207). Some of the genes involved in starch biosynthesis have been cloned from cassava, including starch branching enzyme I (SBE I) (Salehuzzaman et al., 1994 Plant Science 98, 53-62), and granule bound starch synthase I (GBSS I) (Salehuzzaman et al., 1993 Plant Molecular Biology 23, 947-962) and some work has been done on their expression patterns although only in in vitro grown plants (Salehuzzaman et al., 1994 Plant Science 98, 53-62).

In most plants studied to date e.g. maize (Boyer & Preiss, 1978 Biochem. Biophys. Res. Comm. 80, 169-175), rice (Smyth. 1988 Plant Sci. 57, 1-8) and pea (Smith, Planta 175, 270-279), two forms of SBE have been identified, each encoded by a separate gene. A recent review by Burton et al.. (1995 The Plant Journal 7, 3-15) has demonstrated that the two forms of SBE constitute distinct classes of the enzyme such that, in general, enzymes of the same class from different plants may exhibit greater similarity than enzymes of different classes from the same plant. In their review, Burton et al. termed the two respective enzyme families class "A" and class "B", and the reader is referred thereto (and to the references cited therein) for a detailed discussion of the distinctions between the two classes. One general distinction of note would appear to be the presence, in class A SBE molecules, of a flexible N-terminal domain, which is not found in class B molecules. The distinctions noted by Burton et al. are relied on herein to define class A and class B SBE

molecules, which terms are to be interpreted accordingly.

Many organisations have interests in obtaining modified Cassava starches by means of genetic modification. This is impossible to achieve however, unless the plant is amenable to transformation and regeneration, and the starch biosynthesis genes which are to be targeted for modification must be cloned. The production of transgenic cassava plants has only recently been demonstrated (Taylor et al., 1996 Nature Biotechnology 14, 726-730; Schöpke et al., 1996 Nature Biotechnology 14, 731-735; and Li et al., 1996 Nature Biotechnology 14, 736-740). The present invention concerns the identification, cloning and sequencing of a starch biosynthetic gene from Cassava, suitable as a target for genetic manipulation.

Summary of the Invention

In a first aspect the invention provides a nucleic acid sequence encoding a polypeptide having starch branching enzyme (SBE) activity, the polypeptide comprising an effective portion of the amino acid sequences shown in Figure 4 or Figure 13. The nucleic acid is conveniently in substantial isolation, especially in isolation from other naturally associated nucleic acid sequences.

An "effective portion" of the amino acid sequences may be defined as a portion which retains sufficient SBE activity when expressed in E. coli KV832 to complement the branching enzyme mutation therein. The amino acid sequences shown in Figures 4 and 13 include the N terminal transit peptide, which comprises about the first 50 amino acid residues. As those skilled in the art will be well aware, such a transit peptide is not essential for SBE activity. Thus the mature polypeptide, lacking a transit peptide, may be considered as one example of an effective portion of the amino acid sequence shown in Figure 4 or Figure 13.

Other effective portions may be obtained by effecting minor deletions in the amino acid sequence, whilst substantially preserving SBE activity. Comparison with known class A SBE sequences, with the benefit of the disclosure herein, will enable those skilled in the

4

art to identify regions of the polypeptide which are less well conserved and so amenable to minor deletion, or amino acid substitution (particularly, conservative amino acid substitution) whilst substantially preserving SBE activity. Such less well-conserved regions are generally found in the N terminal amino acid residues (up to the triple proline "elbow" at residues 138-140 in Figure 4 and up to the proline elbow at residues 143-145 in Figure 13) and in the last 50 residues or so of the C terminal, and in particular in the acidic tail of the C terminal.

Conveniently the nucleic acid sequence is obtainable from cassava, preferably obtained therefrom, and typically encodes a polypeptide obtainable from cassava. In a particular embodiment, the encoded polypeptide may have the amino acid sequence NSKH at about position 697 (in relation to Figure 4), which sequence appears peculiar to an isoform of the SBE class A enzyme of cassava, other class A SBE enzymes having the conserved sequence DA D/E Y (Burton et al., 1995 cited above).

In a particular aspect of the invention there is provided a nucleic acid comprising a portion of nucleotides 21 to 2531 of the nucleic acid sequence shown in Figure 4, or a functionally equivalent nucleic acid sequence. Such functionally equivalent nucleic acid sequences include, but are not limited to, those sequences which encode substantially the same amino acid sequence but which differ in nucleotide sequence from that shown in Figure 4 by virtue of the degeneracy of the genetic code. For example, a nucleic acid sequence may be altered (e.g. "codon optimised") for expression in a host other than cassava, such that the nucleotide sequence differs substantially whilst the amino acid sequence of the encoded polypeptide is unchanged. Other functionally equivalent nucleic acid sequences are those which will hybridise under stringent hybridisation conditions (e.g. as described by Sambrook et al., Molecular Cloning. A Laboratory Manual, CSH, i.e. washing with 0.1xSSC, 0.5% SDS at 68°C) with the sequence shown in Figure 4. Figure 10 shows a functionally equivalent sequence designated "125 + 94", which includes a region corresponding to the 3' coding portion of the sequence in Figure 4. Figure 13 shows a functionally equivalent sequence which comprises a second complete SBE coding sequence (the SBE-derived sequence is from nucleotides 35 to 2760, of which the coding sequence is nucleotides 131-2677, the rest of the sequence in the figure is vector-derived).

Functionally equivalent DNA sequences will preferably comprise at least 200-300bp, more preferably 300-600bp, and will exhibit at least 88% identity (more preferably at least 90%, and most preferably at least 95% identity) with the corresponding region of the DNA sequence shown in figures 4 or 10. Those skilled in the art will readily be able to conduct a sequence alignment between the putative functionally equivalent sequence and those detailed in Figures 4 or 10 - the identity of the two sequences is to be compared in those regions which are aligned by standard computer software, which aligns corresponding regions of the sequences.

In particular embodiments the nucleic acid sequence may alternatively comprise a 5' and/or a 3' untranslated region ("UTR"), examples of which are shown in Figures 2 and 4. Figure 9 includes a 3' UTR, as nucleotides 688-1044 and Figure 10 includes 3' UTR as nucleotides 1507-1900 (which nucleotides correspond to the first base after the "stop" codon to the base immediately preceding the poly (A) tail). Any one of the sequences defined above, or a functional equivalent thereof (as defined by hybridisation properties, as set out in the preceding paragraph), could be useful in sense or anti-sense inhibition of corresponding genes, as will be apparent to those skilled in the art. It will also be apparent to those skilled in the art that such regions may be modified so as to optimise expression in a particular type of host cell and that the 5' and/or 3' UTRs could be used in isolation, or in combination with a coding portion of the sequence of the invention. Similarly, a coding portion could be used without a 5' or a 3' UTR if desired.

In a further aspect, the invention provides a replicable nucleic acid construct comprising any one of the nucleic acid sequences defined above. The construct will typically comprise a selectable marker and may allow for expression of the nucleic acid sequence of the invention. Conveniently the vector will comprise a promoter (especially a promoter sequence operable in a plant and/or a promoter operable in a bacterial cell) and one or more regulatory signals known to those skilled in the art.

In another aspect the invention provides a polypeptide having SBE activity, the polypeptide comprising an effective portion of the amino acid sequence shown in Figure 4 or Figure 13. The polypeptide is conveniently one obtainable from cassava, although it may be

derived using recombinant DNA techniques. The polypeptide is preferably in substantial isolation from other polypeptides of plant origin, and more preferably in substantial isolation from any other polypeptides. The polypeptide may have amino acid residues NSKH at about position 697 (in the sequence shown in Figure 4), instead of the sequence DA D/E Y found in other SBE class A polypeptides. The polypeptide may be used in a method of modifying starch in vitro, the method comprising treating starch under suitable conditions (of temperature, pH etc.) with an effective amount of the polypeptide.

Those skilled in the art will appreciate that the disclosure of the present specification can be utilised in a number of ways. In particular, the characteristics of a host cell may be altered by recombinant DNA techniques. Thus, in a further aspect, there is provided a method by which a host cell may be altered by introduction of a nucleic acid sequence comprising at least 200bp and exhibiting at least 88% sequence identity (more preferably at least 90%, and most preferably at least 95% identity) with the corresponding region of the DNA sequence shown in Figures 4, 9, 10 or 13, operably linked in the sense or (preferably) in the anti-sense orientation to a suitable promoter active in the host cell, and causing transcription of the introduced nucleic acid sequence, said transcript and/or the translation product thereof being sufficient to interfere with the expression of a homologous gene naturally present in said host cell, which homologous gene encodes a polypeptide having SBE activity. The altered host cell is typically a plant cell, such as a cell of a cassava, banana, potato, sweet potato, tomato, pea, wheat, barley, oat, maize, or rice plant.

Desirably the method further comprises the introduction of one or more nucleic acid sequences which are effective in interfering with the expression of other homologous gene or genes naturally present in the host cell. Such other genes whose expression is inhibited may be involved in starch biosynthesis (e.g. an SBE I gene), or may be unrelated to SBE II.

Those skilled in the art will be aware that both anti-sense inhibition, and "sense suppression" of expression of genes, especially plant genes, has been demonstrated (e.g. Matzke & Matzke 1995 Plant Physiol. 107, 679-685).

It is believed that antisense methods are mainly operable by the production of antisense mRNA which hybridises to the sense mRNA, preventing its translation into functional polypeptide, possibly by causing the hybrid RNA to be degraded (e.g. Sheehy et al., 1988 PNAS 85, 8805-8809; Van der Krol et al., Mol. Gen. Genet. 220, 204-212). Sense suppression also requires homology between the introduced sequence and the target gene, but the exact mechanism is unclear. It is apparent however that, in relation to both antisense and sense suppression, neither a full length nucleotide sequence, nor a "native" sequence is essential. Preferably the nucleic acid sequence used in the method will comprise at least 200-300bp, more preferably at least 300-600bp, of the full length sequence, but by simple trial and error other fragments (smaller or larger) may be found which are functional in altering the characteristics of the plant. It is also known that untranslated portions of sequence can suffice to inhibit expression of the homologous gene - coding portions may be present within the introduced sequence, but they do not appear to be essential under all circumstances.

The inventors have discovered that there are at least two class A SBE genes in cassava. A fragment of a second gene has been isolated, which fragment directs the expression of the C terminal 481 amino acids of cassava class A SBE (see Figure 10) and comprises a 3' untranslated region. Subsequently, a complete clone of the second gene was also recovered (see Figure 12). The coding portions of the two genes show some slight differences, and the second SBE gene may be considered as functionally equivalent to the corresponding portion of the nucleotide sequence shown in Figure 4. However, the 3' untranslated regions of the two genes show marked differences. Thus the method of altering a host cell may comprise the use of a sufficient portion of either gene so as to inhibit the expression of the naturally occurring homologous gene. Conveniently, a portion of nucleotide sequence is employed which is conserved between both genes. Alternatively, sufficient portions of both genes may be employed, typically using a single construct to direct the transcription of both introduced sequences.

In addition, as explained above, it may be desired to cause inhibition of expression of the class B SBE (i.e. SBE I) in the same host cell. A number of class B SBE gene sequences are known, including portions of the cassava class B SBE (Salehuzzaman et al., 1994)

Plant Science 98, 53-62) and any one of these may prove suitable. Preferably the sequence used is that which derives from the host cell sought to be altered (e.g. when altering the characteristics of a cassava plant cell, it is generally preferred to use sense or anti-sense sequences corresponding exactly to at least portions of the cassava gene whose expression is sought to be inhibited).

In a further aspect the invention provides an altered host cell, into which has been introduced a nucleic acid sequence comprising at least 200bp and exhibiting at least 88% sequence identity (more preferably at least 90%, and most preferably at least 95% identity) with the corresponding region of the DNA sequence shown in Figures 4, 9, 10 or 13, operably linked in the sense or anti-sense orientation to a suitable promoter, said host cell comprising a natural gene sharing sequence homology with the introduced sequence.

The host cell may be a micro-organism (such as a bacterial, fungal or yeast cell) or a plant cell. Conveniently the host cell altered by the method is a cell of a cassava plant, or another plant with starch storage reserves, such as banana, potato, sweet potato, tomato, pea, wheat, barley, oat, maize, or rice plant. Typically the sequence will be introduced in a nucleic acid construct, by way of transformation, transduction, micro-injection or other method known to those skilled in the art. The invention also provides for a plant into which has been introduced a nucleic acid sequence of the invention, or the progeny of such a plant.

The altered plant cell will preferably be grown into an altered plant, using techniques of plant growth and cultivation well-known to those skilled in the art of re-generating plantlets from plant cells.

The invention also provides a method of obtaining starch from an altered plant, the plant being obtained by the method defined above. Starch may be extracted from the plant by any of the known techniques (e.g. milling). The invention further provides starch obtainable from a plant altered by the method defined above, the starch having altered properties compared to starch extracted from an equivalent but unaltered plant. Conveniently the altered starch is obtained from an altered plant selected from the group

consisting of cassava, potato, pea, tomato, maize, wheat, barley, oat, sweet potato and rice. Typically the altered starch will have increased amylose content.

The invention will now be further described by way of illustrative examples and with reference to the accompanying drawings, in which:-

Figure 1 is a schematic illustration of the cloning strategy for cassava SBE II. The top line represents the size of a full length clone with distances in kilobases (kb) and arrows representing oligonucleotides (rightward pointing arrows are sense strand, leftward are on opposite strand). The long thick arrow is the open reading frame with start and stop codons shown. Below this are shown the 3' RACE, 5' RACE and PCR clones identified either by the plasmid name (shown in brackets above the line) or the clone number (shown to the left of the clone) for the 5' RACE only. Also shown (by an x) in the 5' RACE clones are positions of small deletions or introns.

Figure 2 shows the DNA sequence and predicted ORF of csbe2con.seq. This sequence is a consensus of 3' RACE pSJ94 and 5' RACE clones 27/9,11 and 28. The first 64 base pairs are derived from the RoRidT17 adaptor primer/dT tail followed by the SBE sequence. The one long open reading frame is shown in one letter code below the double strand DNA sequence. Also shown is the upstream ORF (MQL...LPW).

Figure 3 shows an alignment of the 5' region of cassava SBE II csbe2con and pSJ99 (clones 20 and 35) DNA sequences. Differences from the consensus sequence are shaded.

Figure 4 shows the DNA sequence and predicted ORF of full length cassava SBE II tuber cDNA in pSJ107. The sequence shown is from the CSBE214 to the CSBE218 oligonucleotide. The DNA sequence is sequence ID No. 28 in the attached sequence listing; the amino acid sequence is Seq ID No. 29.

Figure 5 shows an alignment of 3' region of cassava SBE II pSJ116 and 125+94 DNA sequences. The top line is the 125 + 94 sequence and the bottom SJ116 sequence. Identical nucleotides are indicated by the same letter in the middle line, differences are

indicated by a gap, and dashed lines indicate gaps introduced to optimise alignment.

Figure 6 shows an alignment of carboxy terminal region of pSJ116 and 125+94 protein sequences. The top sequence is from 125+94 and the bottom from pSJ116. Identical amino acid residues are shown with the same letter, conserved changes with a colon and neutral changes with a period.

Figure 7 shows a phylogenetic tree of starch branching enzyme proteins. The length of each pair of branches represents the distance between sequence pairs. The scale beneath the tree measures the distance between sequences (units indicate the number of substitution events). Dotted lines indicate a negative branch length because of averaging the tree. Zmcon12.pro is maize SBE II. psstb1.pro is pea SBE I (Bhattacharyya et al 1990 Cell 60, 115-121) and atsbe2-1 & 2-2.pro are two SBE II proteins from Arabidopsis thalania (Fisher et al 1996 Plant Mol. Biol. 30, 97-108). SJ107.pro is representative of a cassava SBE II sequence, and potsbe2.pro is a potato SBE II sequence known to the inventors.

Figure 8 is an alignment of SBE II proteins. Protein sequences are indicated in one letter code. The top line represents the consensus sequence, below which is shown the consensus ruler and the individual SBE II sequences. Residues matching the consensus are shaded. Dashes represent gaps introduced to optimise alignment. Sequence identities are shown at the right of the figure and are as Figure 7, except that SJ107.pro is cassava SBE II.

Figure 9 shows the DNA sequence and predicted ORF of a cassava SBE II cDNA isolated by 3' RACE (plasmid pSJ 101).

Figure 10 shows the consensus DNA sequence and predicted ORF of a second cassava SBE II cDNA isolated by 3' and 5' RACE (sequence designated 125+94 is from plasmid pSJ125 and pSJ94, spliced at the CSBE217 oligo sequence).

Figure 11 is a schematic diagram of the plant transformation vector pSJ64. The black line represents the DNA sequence. The hashed line represents the bacterial plasmid backbone

(containing the origin of replication and bacterial selection marker) and is not shown in full. The filled triangles represent the T-DNA borders (RB = right border, LB = left border). Relevant restriction enzyme sites are shown above the black line with the approximate distances (in kiloobases) betwen sites marked by an asterisk shown underneath. The thinnest arrows represent polyadenylation signals (pAnox = nopaline synthase, pAg7 = Agrobacterium gene 7), the intermediate arrows represent protein coding regions (SBE II = cassava SBE II, HYG = hygromycin resistance gene) and the thick arrows represent promoter regions (P-2x35S = double CaMV 35S promoter, P-nos = nopaline synthase promoter).

Figure 12 is a schematic illustration of the cloning strategy used to isolate a second cassava SBE II gene. The top line represents the size of a full length clone with distances in kilobases (kb) and arrows representing oligonucleotides (rightward pointing arrows are sense strand, leftward are on opposite strand). The long thick arrow is the open reading frame with start and stop codons shown. Below this are shown the 3' RACE, 5' RACE and PCR clones identified either by the plasmid name (shown in brackets above the line) or the clone number (shown to the right of the clone).

Figure 13 shows the DNA sequence and predicted ORF of a second full length cassava SBE II tuber cDNA in pSJ146. Nucleotides 35-2760 are SBE II sequence and the remainder are from the pT7Blue vector. The DNA sequence of Figure 13 is Seq ID No. 30, and the amino acid sequence is Seq ID No. 31, in the attached sequence listing.

Example 1

This example relates to the isolation and cloning of SBE 11 sequences from cassava.

Recombinant DNA manipulations

Standard procedures were performed essentially according to Sambrook et al. (1989 Molecular cloning A laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.). DNA sequencing was performed on an ABI automated DNA sequencer and sequences manipulated using DNASTAR software for the Macintosh.

Rapid Amplification of cDNA ends (RACE) and PCR conditions

5' and 3' RACE were performed essentially according to Frohman et al., (1988 Proc. Natl. Acad. Sci. USA 85, 8998-9002) but with the following modifications.

For 3' RACE, 5 μ g of total RNA was reverse transcribed using 5 pmol of the RACE adaptor RoRidT17 as primer and Stratascript RNAse H- reverse transcriptase (50 U) in a 50 μ l reaction according to the manufacturer's instructions (Stratagene). The reaction was incubated for 1 hour at 37°C and then diluted to 200 μ l with TE (10 mM Tris HCl, 1 mM EDTA) pH 8 and stored at 4°C. 2.5 μ l of this cDNA was used in a 25 μ l PCR reaction with 12.5 pmol of SBE A and Ro primers for 30 cycles of 94°C 45 sec, 50°C 25 sec, 72°C 1 min 30 sec. A second round of PCR (25 cycles) was performed using 1 μ l of this reaction as template in a 50 μ l reaction under the same conditions. Amplified products were separated by agarose gel electrophoresis and cloned into the pT7Blue vector (Invitrogen).

For the first round of 5° RACE, 5 µg of total leaf RNA was reverse transcribed as described above using 10 pmol of the SBE II gene specific primer CSBE22. This primer was removed from the reaction by diluting to 500 µl with TE and centrifuging twice through a centricon 100 microconcentrator. The concentrated cDNA was then dA-tailed with 9U of terminal deoxynucleotide transferase and 50 µM dATP in a 20 µl reaction in huffer supplied by the manufacturer (BRL). The reaction was incubated for 10 min at 37°C and 5 min at 65°C and then diluted to 200 µl with TE pH 8. PCR was performed in a 50 µl volume using 5µl of tailed cDNA, 2.5 pmol of RoRidT17 and 25 pmol of Ro and CSBE24 primers for 30 cycles of 94°C 45 sec, 55°C 25 sec, 72°C 3 min. Amplified products were separated on a 1% TAE agarose gel, cut out, 200µl of TE was added and melted at 99°C for 10 min. Five µl of this was re-amplified in a 50 µl volume using CSBE25 and Ri as primers and 25 cycles of 94°C 45 sec. 55°C 25 sec, 72°C 1 min 30 sec. Amplified fragments were separated on a 1% TAE agarose gel, purified on DEAE paper and cloned into pT7Blue.

The second round of 5' RACE was performed using CSBE28 and 29 primers in the first and second round PCR reactions respectively using a new A-tailed cDNA library primed

with CSBE27.

A third round of 5' RACE was performed on the same CSBE27 primed cDNA.

Repeat 3' RACE and PCR Cloning

The 3' RACE library (RoRidT17 primed leaf RNA) was used as a template. The first PCR reaction was diluted 1:20 and 1 μ l was used in a 50 μ l PCR reaction with SBE A and Ri primers and the products were closed into pT7Blue. The closed PCR products were screened for the presence or absence of the CSBE23 oligo by colony PCR.

A full length cDNA of cassava SBE II was isolated by PCR from leaf or root cDNA (RoRidT17 primed) using primers CSBE214 and CSBE218 from 2.5 μ I of cDNA in a 25 μ I reaction and 30 cycles of 94°C 45 sec, 55°C 25 sec, 72°C 2 min.

Complementation of E. coli mutant KV832

SBE II containing plasmids were transformed into the branching enzyme deficient mutant E. coli KV832 (Keil et al., 1987 Mol. Gen. Genet. 207, 294-301) and cells grown on solid PYG media (0.85 % KH₂PO₄, 1.1 % K₂HPO₄, 0.6 % yeast extract) containing 1.0 % glucuse. To test for complementation, a loop of cells was scraped off and resuspended in 150 μ L water to which was added 15 μ L of Lugol's solution (2 g KI and 1 g I_2 per 300 ml water).

RNA isolation

RNA was isolated from cassava plants by the method of Logemann (1987 Anal. Biochem. 163, 21-26). Leaf RNA was isolated from 0.5 gm of in vitro grown plant tissue. The total yield was 300 μ g. Three month old roots (88 gm) were used for isolation of root RNA).

SBE II specific oligonucleotides

SBE A ATGGACAAGGATATGTATGA (Seq ID No. 1)

CSBE21 GGTTTCATGACTTCTGAGCA (Seq ID No. 2)

CSBE22	TGCTCAGAAGTCATGAAACC	(Seq ID No. 3)
CSBE23	TCCAGTCTCAATATACGTCG	(Seq ID No. 4)
CSBE24	AGGAGTAGATGGTCTGTCGA	(Seq ID No. 5)
CSBE25	TCATACATATCCTTGTCCAT	(Seq ID No. 6)
CSBE26	GGGTGACTTCAATGATGTAC	(Seq ID No. 7)
CSBE27	GGTGTACATCATTGAAGTCA	(Seq ID No. 8)
CSBE28	AATTACTGGCTCCGTACTAC	(Seq ID No. 9)
CSBE29	CATTCCAACGTGCGACTCAT	(Seq ID No. 10)
CSBE210	TACCGGTAATCTAGGTGTTG	(Seq ID No. 11)
CSBE211	GGACCTTGGTTTAGATCCAA	(Seq ID No. 12)
CSBE212	ATGAGTCGCACGTTGGAATG	(Seq ID No. 13)
CSBE213	CAACACCTAGATTACCGGTA	(Seq ID No. 14)
CSBE214	TTAGTTGCGTCAGTTCTCAC	(Seq ID No. 15)
CSBE215	AATATCTATCTCAGCCGGAG	(Seq ID No. 16)
CSBE216	ATCTTAGATAGTCTGCATCA	(Seq ID No. 17)
CSBE217	TGGTTGTTCCCTGGAATTAC	(Seq ID No. 18)
CSBE218	TGCAAGGACCGTGACATCAA	(Seq 1D No. 19)

RESULTS

Cloning of a SBE II gene from cassava leaf

The strategy for cloning a full length cDNA of starch branching enzyme II of cassava is shown in Figure 1. A comparison of several SBE II (class A) SBE DNA sequences identified a 23 bp region which appears to be completely conserved among most genes (data not shown) and is positioned about one kilobase upstream from the 3' end of the gene. An oligonucleotide primer (designated SBE A) was made to this sequence and used to isolate a partial cDNA clone by 3' RACE PCR from first strand leaf cDNA as illustrated in Figure 1. An approximately 1100 bp hand was amplified, cloned into pT7Blue vector and sequenced. This clone was designated pSJ94 and contained a 1120 bp insert starting with the SBE A oligo and ending with a polyA tail. There was a predicted open reading frame of 235 amino acids which was highly homologous (79% identical) to a potato SBE II also isolated by the inventors (data not shown) suggesting that this clone represented a class A (SBE II) gene.

To obtain the sequence of a full length clone nested primers were made complementary to the 5' end of this sequence and used in 5' RACE PCR to isolate clones from the 5' region of the gene. A total of three rounds of 5' RACE was needed to determine the sequence of the complete gene (i.e. one that has a predicted long ORF preceded by stop codons). It should be noted that during this cloning process several clones (#23, 9, 16) were obtained that had small deletions and in one case (clone 23) there was also a small (120 bp) intron present. These occurrences are not uncommon and probably arise through errors in the PCR process and/or reverse transcription of incompletely processed RNA (heterogeneous nuclear RNA).

The overlapping cDNA fragments could be assembled into a contiguous 3 kb sequence (designated csbe2con.seq) which contained one long predicted ORF as shown in Figure 2. Several clones in the last round of 5' RACE were obtained which included sequence of the untranslated leader (UTL). All of these clones had an ORF (42 amino acids) 46 bp upstream and out of frame with that of the long ORF.

There is more than one SBE II gene in cassava

In order to determine if the assembled sequence represented that of a single gene, attempts were made to recover by PCR a full length SBE II gene using primers CSBE214 and CSBE23 at the 5' and 3' ends of the csbe2con sequence respectively. All attempts were unsuccessful using either leaf or root cDNA as template. The PCR was therefore repeated with either the 5'- or 3'- most primer and complementary primers along the length of the SBE II gene to determine the size of the largest fragment that could be amplified. With the CSBE214 primer, fragments could be amplified using primers 210, 28, 27 and 22 in order of increasing distance, the latter primer pair amplifying a 2.2 kb band. With the 3' primer CSBE23, only primer pairs with 21 and 26 gave amplification products, the latter being about 1200 bp. These results suggest that the original 3' RACE clone (pSJ94) is derived from a different SBE II gene than the rest of the 5' RACE clones even though the two largest PCR fragments (214+22 and 26+23) overlap by 750 bp and share several primer sites. It is likely that the sequence of the two genes starts to diverge around the CSBE22 primer site such that the 3' end of the corresponding gene does not contain the 23 primer and is not therefore able to amplify a cDNA when used with the 214 primer.

To confirm this, the sequence of the longest 5' PCR fragment (214+22) from two clones (#20 designated pSJ99. & #35) was determined and compared to the consensus sequence csbe2con as shown in Figure 3. The first 2000 bases are nearly identical (the single base changes might well be PCR errors), however the consensus sequence is significantly different after this. This region corresponds to the original 3' RACE fragment pSJ94 (SBE A + Ri adaptor) and provided evidence that there may be more than one SBE II gene in cassava.

The 3' end corresponding to pSJ99 was therefore cloned as follows: 3' RACE PCR was performed on leaf cDNA using the SBE A oligo as the gene specific primer so that all SBE II genes would be amplified. The cloned DNA fragments were then screened for the presence or absence of the CSBE23 primer by PCR. Two out of 15 clones were positive with the SBE A + Ri primer pair but negative with SBE A + CSBE23 primers. The sequence of these two clones (designated pSJ101, as shown in Figure 9) demonstrated that they were indeed from an SBE II gene and that they were different from pSJ94. However the overlapping region of pSJ101 (the 3' clone) and pSJ99 (the 5' clone) was identical suggesting that they were derived from the same gene.

To confirm this a primer (CSBE218) was made to a region in the 3' UTR (untranslated region) of pSJ101 and used in combination with CSBE214 primer to recover by PCR a full length cDNA from both leaf and root cDNA. These clones were sequenced and designated pSJ106 & pSJ107 respectively. The sequence and predicted ORF of pSJ107 is shown in Figure 4. The long ORF in plasmid pSJ106 was found to be interrupted by a stop codon (presumably introduced in the PCR process) approximately 1 kb from the 3' end of the gene, therefore another cDNA clone (designated pSJ116) was amplified in a separate reaction, cloned and sequenced. This clone had an intact ORF (data not shown). There were only a few differences in these two sequences (in the transit peptide as 27-41: YRRTSSCLSFNFKEA to DRRTSSCLSFIFKKAA and L831 in pSJ107 to V in pSJ116 respectively).

An additional 740bp of sequence of the gene corresponding to the pSJ94 clone was isolated by 5' RACE using the primers CSBE216 and 217, and was designated pSJ125.

PCT/GB97/03032

This sequence was combined with that of pSJ94 to form a consensus sequence "125 + 94", as shown in Figure 10. The sequence of this second gene is about 90% identical at the DNA and protein level to pSJ116, as shown in Figure 5 and 6, and is clearly a second form of SBE II in cassava. The 3' untranslated regions of the two genes are not related (data not shown).

17

It was also determined that the full length cassava SBE II genes (from both leaf and tuber) actually encode for active starch branching enzymes since the cloned genes were able to complement the glycogen branching enzyme deficient E. coli mutant KVR32.

Main Findings

- 1) A full length cDNA clone of a starch branching enzyme II (SBE II) gene has been cloned from leaves and starch storing mots of cassava. This cDNA encodes a 836 amino acid protein (Mr 95 Kd) and is 86 % identical to pea SBE I over the central conserved domain, although the level of sequence identity over the entire coding region is lower than 86%.
- 2) There is more than one SBE II gene in cassava as a second partial SBE II cDNA was isolated which differs slightly in the protein coding region from the first gene and has no homology in the 3' untranslated region.
- 3) The isolated full length cDNA from both leaves and roots encodes an active SBE as it complements an E. coli mutant deficient in glycogen branching enzyme as assayed by iodine staining.

We have shown that there are SBE II (Class A) gene sequences present in the cassava genome by isolating cDNA fragments using 3' and 5' RACE. From these cDNA fragments a consensus sequence of over 3 kb could be compiled which contained one long open reading frame (Figure 2) which is highly homologous to other SBE II (class A) genes (data not shown). It is likely that the consensus sequence does not represent that of a single gene since attempts to PCR a full length gene using primers at the 5' and 3' ends of this sequence were not successful. In fact screening of a number of leaf derived 3'

RACE cDNAs showed that a second SBE 11 gene (clone designated pSJ101) was also expressed which is highly homologous within the coding region to the originally isolated cDNA (pSJ94) but has a different 3' UTR. A full length SBE II gene was isolated from leaves and roots by PCR using a new primer to the 3' end of this sequence and the original sequence at the 5' end of the consensus sequence. If the frequency of clones isolated by 3' RACE PCR reflects the abundance of the mRNA levels then this full length gene may be expressed at lower levels in the leaf than the pSJ94 clone (2 out of 15 were the former class, 13/15 the latter). It should be noted that each class is expressed in both leaves and roots as judged by PCR (data not shown). Sequence analysis of the predicted ORF of the leaf and root genes showed only a few differences (4 amino acid changes and one deletion) which could have arisen through PCR errors or, alternatively, there may be more than one nearly identical gene expressed in these tissues.

A comparison of all known SBE II protein sequences shows that the cassava SBE II gene is most closely related to the pea gene (Figure 8). The two proteins are 86.3% identical over a 686 amino acid range which extends from the triple proline "elbow" (Burton et al., 1995 Plant J. 7, 3-15) to the conserved VVYA sequence immediately preceding the Cterminal extensions (data not shown). All SBE II proteins are conserved over this range in that they are at least 80% similar to each other. Remarkably however, the sequence conservation between the pea, potato and cassava SBE II proteins also extends to the Nterminal transit peptide, especially the first 12 amino acids of the precursor protein and the region surrounding the mature terminus of the pea protein (AKFSRDS). Because the proteins are so similar around this region it can be predicted that the mature terminus of the cassava SBE II protein is likely to be GKSSHES. The precursor has a predicted molecular mass of 96 kD and the mature protein a predicted molecule mass of 91.3 kD. The cassava SBE II has a short acidic tail at the C-terminal although this is not as long or as acidic as that found in the pea or potato proteins. The significance of this acidic tail, if any, remains to be determined. One notable difference between the amino acid sequence of cassava SBE II and all other SBE II proteins is the presence of the sequence NSKH at around position 697 instead of the conserved sequence DAD/EY. Although this conserved region forms part of a predicted a-helix (number 8) of the catalytic $(B/a)_a$ barrel domain (Burton et al 1995 cited previously), this difference does not abolish the SBE activity of the cassava protein as this gene can still complement the glycogen branching deletion mutant of E. coli. It may however affect the specificity of the protein. An interesting point is that the other cassava SBE II clone pSJ94 has the conserved sequence DADY.

One other point of interest concerning the sequence of the SBE II gene is the presence of an upstream ATG in the 5° UTR. This ATG could initiate a small peptide of 42 amino acids which would terminate downstream of the predicted initiating methionine codon of the SBE II precursor. If this does occur then the translation of the SBE II protein from this mRNA is likely to be inefficient as ribosomes normally initiate at the 5° most ATG in the mRNA. However the first ATG is in a poorer Kozak context than the SBE II initiator and it may be too close to the 5° end of the message to initiate efficiently (14 nucleotides) thus allowing initiation to occur at the correct ATG.

In conclusion we have shown that cassava does have SBE II gene sequences, that they are expressed in both leaves and tubers and that more than one gene exists.

Example 2

Cloning of a second full length cassava SBE II gene

Methods

Olizonucleotides

CSBE219	CTTTATCTATTAAAGACTTC	(Seq ID No. 20)
CSBE220	CAAAAAGTTTGTGACATGG	(Seq ID No. 21)
CSBE221	TCACTTTTTCCAATGCTAAT	(Seq ID No. 22)
CSBE222	TCTCATGCAATGGAACCGAC	(Seq ID No. 23)
CSBE223	CAGATGTCCTGACTCGGAAT	(Seq ID No. 24)
CSBE224	ATTCCGAGTCAGGACATCTG	(Seq ID No. 25)
CSBE225	COCATITCTCGCTATTGCTT	(Seq ID No. 26)
CSBE226	CACAGGCCCAAGTGAAGAAT	(Seq ID No. 27)

The 5' end of the gene corresponding to the 3'RACE clone pSJ94 was isolated in three

rounds of 5'RACE. Prior to performing the first round of 5' RACE, 5 μ g of total leaf RNA was reverse transcribed in a 20 μ l reaction using conditions as decribed by the manufacturer (Superscript enzyme, BRL) and 10 pmol of the SBE II gene specific primer CSBE23. Primers were then removed and the cDNA tailed with dATP as described above. The first round of 5'RACE used primers CSBE216 and Ro. This PCR reaction was diluted 1:20 and used as a template for a second round of amplification using primers CSBE217 and Ri. The gene specific primers were designed so that they would preferentially hybridise to the SBE II sequence in pSJ94. Amplified products appeared as a smear of approximately 600-1200 bp when subjected to electrophoresis on a 1% TAE agarose gel.

This smear was excised and DNA purified using a Qiaquick column (Qiagen) before ligation to the pT7Blue vector. Several clones were sequenced and clone #7 was designated pSJ125. New primers (CSBE219 and 220) were designed to hybridise to the 5' end of pSJ125 and a second round of 5'RACE was performed using the same CSBE23 primed library. Two fragments of 600 and 800 bp were cloned and sequenced (clones 13,17). Primers CSBE221 and 222 were designed to hybridise to the 5' sequence of the longest clone (#13) and a third round of 5' RACE was performed on a new library (5 μ g total leaf RNA reverse transcribed with Superscript using CSBE220 as primer and then dATP tailed with TdT from Boehringer Mannheim). Fragments of approximately 500 bp were amplified, cloned and sequenced. Clone #13, was designated pSJ143. The process is illustrated schematically in Figure 12.

To isolate a full length gene as a contiguous sequence, a new primer (CSBE225) was designed to hybridise to the 5' end of clone pSJ143 and used with one of the primers (CSBE226 or 23) in the 3' end of clone pSJ94, in a PCR reaction using RoRidT17 primed leaf cDNA as template. Use of primer CSBE226 resulted in production of Clone #2 (designated pSJ144), and use of primer CSBE23 resulted in production of Clones #10 and 13 (designated pSJ145 and pSJ146 respectively). Only pSJ146 was sequenced fully.

Results

Isolation of a second full length cassava SBE II gene

A full length clone for a second SBE II gene was isolated by extending the sequence of pSJ94 in three rounds of 5' RACE as illustrated schematically in Figure 12. In each round of 5' RACE, primers were designed that would preferentially hybridise to the new sequence rather than to the gene represented by pSJ116. In the final round of 5' RACE, three clones were obtained that had the initiating methione codon, and none of these had upstream ATGs. The overlapping cDNA fragments (sequences of the 5'RACE clones pSJ143, 13, pSJ125 and the 3'RACE clone pSJ94) could be assembled into a consensus sequence of approximately 3 kb which was designated csbe2-2.seq. This sequence contained one long ORF with a predicted size of 848 aa (M, 97 kDa). The full length gene was then isolated as a contiguous sequence by PCR amplification from RoRidT17 primed leaf cDNA using primers at the 5' (CSBE225) and 3' (CSBE23 or CSBE226) ends of the RACE clones. One clone, designated pSJ146, was sequenced and the restriction map is shown along with the predicted amino acid sequence in Figure 13.

Sequence homologies between SBE II genes

The two cassava genes (pSJ116 and pSJ146) share 88.8% identity at the DNA level over the entire coding region (data not shown). The homology extends about 50 bases outside of this region but beyond this the untranslated regions show no similarity (data not shown). At the protein level the two genes show 86% identity over the entire ORF (data not shown). The two genes are more closely related to each other than to any other SBE II. Between species, the pea SBE I shows the most homology to the cassava SBE II genes.

Example 3

Construction of plant transformation vectors and transformation of cassava with authense starch branching enzyme genes.

This example describes in detail how a portion of the SBE II gene isolated from cassava may be introduced into cassava plants to create transgenic plants with altered properties.

An 1100 bp *Hind* III - *Sac* I fragment of cassava SBE II (from plasmid pSJ94) was cloned into the *Hind* III - *Sac* I sites of the plant transformation vector pSJ64 (Figure 11). This placed the SBE II gene in an antisense orientation between the 2X 35S CaMV promoter

- and the nopaline synthase polyadenylation signal. PSJ84 is a derivative o the binary vector pGPTV-HYG (Becker et al., 1992 Plant Molecular Biology 20: 1195-1197) modified by inclusion of an approximately 750 bp fragment of pJIT60 (Guerineau et al 1992 Plant Mol. Biol. 18, 815-818) containing the duplicated cautiflower mosaic virus (CaMV) 35S promoter (Cabb-JI strain, equivalent to nucleotides 7040 to 7376 duplicated upstream of 7040 to 7433, as described by Frank et al., 1980 Cell 21, 285-294) to replace the GUS coding sequence. A similar construct was made with the cassava SBE II sequence from plasmid pSJ101.
- 1.5 These plasmids are then introduced into Agrobacterium tumefaciens LBA4404 by a direct DNA uptake method (An et al. Binary vectors, In: Plant Molecular Biology Manual (ed embryos by selecting on hygromycin as described by Li et al. (1996, Nature Biotechnology 14, 738-740).
- 20 The term "comprise" and grammatical variations thereof such as "comprising" when used in the description or claims does not preclude the presence of additional features, integers, steps or components.

25

30



23

SEQUENCE LISTING

	•	
(1) (GENERAL INFORMATION:	
	(1) APPLICANT: (A) NAME: National Starch and Chemical Investment Holding Corporation (B) STREET: Suite 27, 501 Silverside Road (C) CITY: Wilmington (D) STATE: Delaware (E) COUNTRY: USA (F) POSTAL CODE (ZIP): 19809	
((ii) TITLE OF INVENTION: Improvements in or Relating to Starch Content of Plants	
. (111) NUMBER OF SEQUENCES: 31 -	
1	(iv) COMPUTER READABLE FORM: (A) MEDIUM TYPE: Floppy disk (B) COMPUTER: IBM PC compatible (C) OPERATING SYSTEM: PC-DOS/MS-DOS (D) SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)	
(2)	INFORMATION FOR SEQ ID NO: 1:	
	(1) SECUENCE CHARACTERISTICS: (A) LENGTH. 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:	
ATGG	ACAAGG ATATGTATGA	2
(2)	INFORMATION FOR SEQ ID NO: 2:	••
•	(1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
	(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 2:	
GGTT	TCATGA CTTCTGAGCA	. 2
(2)	INFORMATION FOR SEQ ID NO: 3:	

SUBSTITUTE SHEET (RULE 26)

(i) SEQUENCE CHARACTERISTICS:

W 98/20145		•	PCT/GB97/03032 -
•	. 24		• .
(B) TYPE: n	20 base pairs nucleic acid IDNESS: single NY: linear	-	
(x1) SEQUENCE DES	CRIPTION: SEQ ID N	10: 3:	_
TGCTCAGAAG TCATGAAACC	-		20
(2) INFORMATION FOR S	SEQ ID NO: 4:		
(B) TYPE: n	: 20 base pairs nucleic acid EDNESS: single		
(x1) SEQUENCE DES	SCRIPTION: SEO ID	10 : 4:	
TCCAGTCTCA ATATACGTCG	3		. 20
(2) INFORMATION FOR S	SEQ ID NO: 5:	-	
(1) SEQUENCE CHA (A) LENGTH: (B) TYPE: r (C) STRANDE (D) TOPOLOG	: 20 base pairs nucleic acid EDNESS: single		
(x1) SEQUENCE DES	SCRIPTION: SEQ ID	NO: 5:	
AGGAGTAGAT GGTCTGTCG	A		20
(2) INFORMATION FOR S	SEQ ID NO: 6:		
(B) TYPE: 1	: 20 base pairs nucleic acid EDNESS: single		
(x1) SEQUENCE DE	SCRIPTION: SEQ ID	NO: 6:	
TCATACATAT CCTTGTCCA	T	•	. 20
(2) INFORMATION FOR	SEQ ID NO: 7:		
(B) TYPE: (C) STRAND	ARACTERISTICS: l: 20 base pairs nucleic acid EDNESS: single IGY: linear	_	

WO 98/20145	PCT/GB97/03032 -
25	• •
(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 7:	
GGGTGACTTC AATGATGTAC	20
(2) INFORMATION FOR SEQ ID NO: 8:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	·•
(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 8:	
GGTGTACATC ATTGAAGTCA	20
(2) INFORMATION FOR SEQ ID NO: 9:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 9:	
AATTACTGGC TCCGTACTAC	20
(2) INFORMATION FOR SEQ ID NO: 10:	
(1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 10:	•
CATTCCAACG TGCGACTCAT	. 50
(2) INFORMATION FOR SEQ ID NO: 11:	
(1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	•
(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 11:	
TACCOSTAAT CTAGGTGTTG.	. 20

PCT/GB97/03632 -										
		٠	٠							
	_			20						

- WO 98/20145

26

(2) INFORMATION FOR SEQ ID NO: 12:	
(1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 12:	
GGACCTTGGT TTAGATCCAA	2
(2) INFORMATION FOR SEQ ID NO: 13:	
(1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TUPOLOGY: linear	
(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 13:	
ATGAGTCGCA CGTTGGAATG	2
(2) INFORMATION FOR SEQ ID NO: 14:	
(1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:	
CAACACCTAG ATTACCGGTA	2
(2) INFORMATION FOR SEQ ID NO: 15:	
(1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid. (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 15:	
TTAGTTGCGT CAGTTCTCAC	;

SUBSTITUTE SHEET (RULE 26) .

(2) INFORMATION FOR SEQ ID NO: 16:

(1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs

WO 98/20145	PC1/GH97/03032 -
. 27	
(B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 16:	<u>.</u>
AATATCTATC TCAGCCGGAG	. 20
(2) 'INFORMATION FOR SEQ ID NO: 17:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 17:	•
ATCTTAGATA GTCTGCATCA	20
(2) INFORMATION FOR SEQ ID NO: 18:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 18:	
TGGTTGTTCC CTGGAATTAC	20
(2) INFORMATION FOR SEQ ID NO: 19:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	•
(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 19:	
TGCAAGGACC GTGACATCAA	20
(2) INFORMATION FOR SEQ ID NO: 20:	٠
IE/ IN UNIVITUAL FOR SEY ID MO. CV.	

SUBSTITUTE SHEET (RULE 26)

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

w	98/20145	PCT/GB97/03632 -
	28	
	(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 20:	•
СТ	TTATCTAT TAAAGACTTC	20
(2)	INFORMATION FOR SEQ ID NO: 21:	
	(1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	•
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:	
CA	AAAAAGTT TGTGACATGG	20
(2)	INFORMATION FOR SEQ ID NO: 22:	
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
	(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 22:	
TC/	ACTITITIC CAATGCTAAT	20
(2)	INFORMATION FOR SEQ ID NO: 23:	
	(1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	:
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:	-
TCT	Catgcaa tggaaccgac	20
(2	INFORMATION FOR SEQ ID NO: 24:	•
	(1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:	
CAC	GATGTCCT GACTCGGAAT	20

- WO 96/20145	PCT/GB97/03032	•
29	• •	
(2) INFORMATION FOR SEQ ID NO: 25:		
(1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear		
(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 25:		
ATTCCGAGTC AGGACATCTG		20
(2) INFORMATION FOR SEQ ID NO: 26:	-	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear		
(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 26:	•	
CGCATTTCTC GCTATTGCTT		20
(2) INFORMATION FOR SEQ ID NO: 27:		
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear		
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:		
CACAGGCCCA AGTGAAGAAT		20
(2) INFORMATION FOR SEQ ID NO: 28:		
(1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2588 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear		
(1x) FEATURE: (A) NAME/KEY: CDS (B) LOCATION:212531		
(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 28:		
CTCTCTAACT TCTCAGCGAA ATG GGA CAC TAC ACC ATA TCA GGA AT Met Gly His Tyr Thr Ile Ser Gly-Il		50

TTT Phe	CCT	TGT Cys	GCT Ala	CCA Pro 15	CTC Leu	TGC Cys	AAÄ Lys	TCT Ser	CAA G1n 20	TCT Ser	ACC Thr	GGC G1y	TTC Phe	CAT H1s 25	GGC G1y	9
TAT Tyr	CGG Arg	AGG Arg	ACC Thr 30	TCC Ser	TCT Ser	TGC Cys	CTT Leu	TCC Ser 35	TTC Phe	AAC Asn	TTC Phe	AAG Lys	GAG Glu 40	GCG Ala	TTT Phe	14
TCT Ser	AGG 'Arg	AGG Arg 45	GTC Val	TTC Phe	TCT Ser	GGA G1y	AAG Lys 50	TCA Ser	TCT Ser	CAT His	GAA G1u	TCT Ser 55	GAC Asp	TCC Ser	TCA Ser	19
aat Asn	GTA Vall 60	ATG Met	GTC Val	ACT Thr	GCT Ala	TCT Ser 65	AAA Lys	AGA Arg	GTC Val	CTT Leu	CCT Pro 70	GAT Asp	GGT Gly	CGG Arg	ATT Ile	24
GAA G1u 75	TGC Cys	TAT Tyr	TCT Ser	TCT Ser	TCA Ser 80	ACA Thr	GAT Asp	CAA G1n	TTG Leu	GAA G1u 85	GCC Ala	CCT Pro	GGC G1y	ACA Thr	GTT Val 90	29
TCA Ser	GAA Glu	GAA Glu	TCC Ser	CAG G1n 95	GTG Val	CTT Leu	ACT Thr	GAT Asp	GTT Val 100	GAG G1u	AGT Ser	CTC Leu	ATT Ile	ATG Met 105	GAT As p	33
GAT Asp	AAG Lys	ATT Ile	GTT Val 110	GAA G1u	GAT Asp	GAA G1u	GTA Val	AAT Asn 115	AAA Lys	GAA G1u	TCT Ser	GTT Va I	CCA Pro 120	ATG Het	CGG Arg	38
GAG G1u	ACA Thr	GTT Val 125	AGC Ser	ATC Ile	AGA Arg	AAA Lys	ATT Ile 130	GGA G1y	TCT Ser	AAA Lys	CCA Pro	AGG Arg 135	TCC Ser	ATT Ile	CCT Pro	43
CCA Pro	CCC Pro 140	GGC G1y	AGA Arg	GGG G1y	CAA G1n	AGA Arg 145	ATA I 1e	TAT Tyr	GAC Asp	ATA Ile	GAT ASP 150	CCA Pro	AGC Ser	TTG Leu	ACA Thr	48
GGC G1y 155	TTT Phe	CGT Arg	CAA Gìn	CAC His	CTA Leu 160	GAT Asp	TAC Tyr	CGG Arg	TAT Tyr	TCA Ser 165	CAG G1n	TAC Tyr	AAA Lys	AGA Arg	CTC Leu 170	530
CGA Arg	GAA Glu	GAA G1u	ATT Ile	GAC ASP 175	AAG Lys	TAT Tyr	GAA G1u	GGT Gly	AGT Ser 180	CTG Leu	GAT Asp	GCA Ala	TTT Phe	TCT Ser 185	CGT Arg	578
GGE Gly	TAT Tyr	GAA G1u	AAG Lys 190	TTT Phe	GGT G1y	TTC Phe	TCA Ser	CGC Arg 195	AGT Ser	GAA G1u	ACA Thr	GGA Gly	ATA Ile 200	ACT Thr	TAT Tyr	620
AGA Arg	GAG G1u	TGG Trp 205	GCA A1a	CCA Pro	GGA G1y	GCT Ala	ACG Thr 210	TGG Trp	GCT Ala	GCA _A1,a	TTG Leu	ATT Ile 215	GGA Gly	GAT Asp	TTC Phe	674
AAT Asn	AAC Asn 220	TGG Trp	AAT Asn	CCT Pro	AAT Asn	GCA Ala 225	GAT Asp	GTC Val	ATG Het	ACT Thr	CAG G1n 230	AAT Asn	GAG G1u	TGT Cys	GGT Gly	72

٧	TC al 35	TGG Trp	GAG G1u	ATC Ile	TTT Phe	TTG Leu 240	CCG Pro	AAT Asn	AAT Asn	GCA Ala	GAT Asp 245	GGT Gly	TCA Ser	CCA Pro	CCA Pro	ATT Ile 250	770
P) (CC	CAT His	GGT G1y	TCT Ser	CGA Arg 255	GTA Val	AAG Lys	ATA Ile	CGC Arg	ATG Met 260	GAT Asp	ACT Thr	CCA Pro	TCT Ser	GGC G1y 265	AAC Asn	818
A	AA .ys	GAT ASP	TCT Ser	ATT Ile 270	CCT Pro	GCT Ala	TGG Trp	ATC I le	AAG Lys 275	TTC Phe	TCA Ser	GTT Val	CAA G1n	GCA Ala 280	CCA Pro	GGT Gly	866
G	AA ilu	CTC Leu -	CCA Pro 285	TAT Tyr	AAT Asn	GGC G1y	ATA I le	TAC Tyr 290	TAT Tyr	GAT Asp	CCT Pro	CCC Pro	GAG G1u 295	GAG Glu	GAG G1u	AAG Lys	914
				AAA Lys													962
1	АТ Уг 115	GAG G1u	TCG Ser	CAC H1s	GTT Val	GGA Gly 320	ATG Met	AGT Ser	AGT Ser	ACG Thr	GAG G1u 325	CCA Pro	GTA Val	ATT	AAC Asn	ACA Thr 330	1010
1	AT yr	GCC Ala	AAC Asn	TTT Phe	AGA Arg 335	GAT Asp	GAT Asp	GTG Va 1	CTT Leu	CCT Pro 340	CGC Arg	ATC Ile	AAA Lys	AAG Lys	CTT Leu 345	GGC G1y	1058
1	AC	AAT Asn	GCT Ala	GTF Val 350	CAG G1n	CTC Leu	ATG Met	GCT Ala	ATT 11e 355	CAA G1n	GAG G1u	CAT His	TCA Ser	TAT Tyr 360	TAT Tyr	GCT Ala	1106
5	GT er	TTT Phe	GGG G1 y 365	TAT Tyr	CAC His	GTC Val	ACA Thr	AAC Asn 370	TTT Phe	TAT Tyr	GCA Ala	GCT Ala	AGC Ser 375	AGC Ser	CGA Arg	TTT Phe	1154
9	GA Sly	ACT Thr 380	CCT Pro	GAT Asp	GAT Asp	TTA Leu	AAG Lys 385	TCT Ser	CTA Leu	ATA Ile	GAT Asp	AAA Lys 390	Ala	CAC His	GAG G1u	TTA Leu	1202
(GT Gly 395	Leu	CTT Leu	GTT Val	CTC Leu	ATG Met 400	GAT Asp	ATT Ile	GTT Val	CAT His	AGC Ser 405	CAT His	GCA Ala	TCA Ser	ACT Thr	AAT Asn 410	1250
1	ACG Thr	TTG Leu	GAT Asp	GGG G1y	CTG Leu 415	Asn	ATG Met	TTT Phe	GAT Asp	GGT G1y 420	Thr	GAT Asp	GGT Gly	CAC His	TAC Tyr 425	Phe	1 29 8
(CAC	TCT Ser	GGA Gly	CCA Pro 430	Arg	GGT	CAT H1s	CAT His	TGG Trp 435	Met	TGG Trp	GAC Asp	TCT Ser	CGC Arg 440	Leu	TTC Phe	1346
1	AAC Asn	TAT	GGG GTy 445	Ser	TGG	GAG Glu	GTT Val	CTA Leu 450	Arg	TTT Phe	CTT Leu	Len	TCA Ser 455	Asn	GCA Ala	AGG Arg	1394

TGG TGG TTG GA Trp Trp Leu As 460	AT GAG TAC AAG sp Glu Tyr Lys 465	Phe Asp Gly	TTC AGA TTT GA Phe Arg Phe As 470	T GGG GTG p Gly Val	1442
ACT TCA ATG AT Thr Ser Met Me 475	IG TAC ACC CAT et Tyr Thr His 480	CAT GGA TTG His Gly Leu	CAG GTA GAT TT Gln Val Asp Ph 485	T ACC GGC e Thr Gly 490	1490
AAC TAC AAT GA Asn Tyr Asn Gl	AA TAC TTT GGA lu Tyr Phe Gly 495	TAT GCA ACT Tyr Ala Thr 500	GAT GTA GAT GC Asp Val Asp Al	T GTG GTT a Val Val 505	1538
Tyr Leu Met Le	TG TTG AAT GAT eu Leu Asn Asp 10	ATG ATT CAT Met Ile His 515	GGT CTC TTC CC Gly Leu Phe Pr 52	o Glu Ala	1586
GTC ACC ATT GO Val Thr Ile GT 525	GT GAA GAT GTI Iy Glu Asp Val	AGT GGA ATG Ser Gly Met 530	CCA ACA GTT TG Pro Thr Val Cy 535	C ATT CCG s Ile Pro	1634
GTT GAA GAT GO Val Glu Asp G1 540	GT GGT GTT GGC ly Gly Val Gly 54	Phe Asp Tyr	CGT CTC CAC AT Arg Leu His Me 550	G GCT GTT t Ala Val	1682
GCT GAT AAA TO Ala Asp Lys To 555	GG GTT GAG AT rp Val Glu Ile 560	ATT CAG AAG Ile Gln Lys	AGA GAT GAA GA Arg Asp Glu As 565	T TGG AAA p Trp Lys 570	1730
ATG GGT GAC AT Met Gly Asp I	TT GTA CAT ATO Te Val His Met 575	CTG ACC AAC Leu Thr Asn 580	AGG CGG TGG TT Arg Arg Trp Le	G GAA AAG u Glu Lys 585	1778
Cys Val Ser Ty	AT GCT GAA AG yr Ala Glu Sei 90	CAT GAC CAG His Asp Gin 595	GCC CTT GTT GG Ala Leu Val Gl	y Asp Lys	1826
ACT ATT GCA T Thr Ile Ala Pi 605	TT TGG CTG ATT he Trp Leu Me	G GAC AAG GAT ASP Lys ASP 610	ATG TAT GAC TT Met Tyr Asp Pr 615	C ATG GCT e Met Ala	1874
CTT GAC AGA CO Leu Asp Arg P 620	CA ICT ACT CC ro Ser Thr Pr 62	<u>p</u> Leu Ile Asp	CGT GGA GTA GC Arg Gly Val Al 630	A TTG CAC a Leu H1s	1922
AAA ATG ATC A Lys Met Ile A 635	GG CTT ATT AC ing Leu Ile Th 640	C ATG GGA TTA r Met Gly Leu	GGC GGA GAA GG GTy GTy GTu_GT 645	A TAT TTG by Tyr Leu 650	1970
AAT TTT ATG G Asn Phe Met G	GA AAT GAA IT ily Asn Glu Ph 655	T GGA CAC CCC e Gly His Pro 660	GAG TGG ATT G Glu Trp Ile As	AT TTT CCA sp Phe Pro 665	2018
Arg Gly Asp L	TA CAT CTT CC eu His Leu Pr 570	C AGT GGT AAA o Ser Gly Lys 675	A TIT GIT CCT GG S Phe Val Pro G 6	G AAC AAT Iy Asn Asn 30	2066

		•														
	AGT Ser															2114
CAT His	CTG Leu 700	AGA Arg	TAT Tyr	CAT H1s	GGA G1y	ATG Net 705	CAA Gìn	GAG Glu	TTT Phe	GAT Asp	CÁA G1n 710	GCA Ala	ATT 11e	CAG G1n	CAT His	2162
	GAA G1u															2210
	AAG Lys															2258
	TTT Phe															2306
	GGC G1y															2354
	CCT Pro 780															2402
	AGC Ser															2450
	ACA Thr															2498
	GAG G1u									TAA *	GAT	ATAT	CTT A	AACA	ACAGGT	2551
TCTGAAGCAG GAATGCCATT ATTGATCTTC CTATGTT								2588								

(2) INFORMATION FOR SEQ ID NO: 29:

- (i) SEQUENCE CHARACTERISTICS; (A) LENGTH: 837 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein (x1) SEQUENCE DESCRIPTION: SEQ 1D NO: 29:

Met Gly His Tyr Thr Ile Ser Gly Ile Arg Phe Pro Cys Ala Pro Leu
15 10 15

Cys Lys Ser Gln Ser Thr Gly Phe His Gly Tyr Arg Arg Thr Ser Ser 20 25 30 Cys Leu Ser Phe Asn Phe Lys Glu Ala Phe Ser Arg Arg Val Phe Ser $\frac{1}{45}$ Gly Lys Ser Ser His Glu Ser Asp Ser Ser Asn Val Met Val Thr Ala $50 \hspace{0.25cm} 55$ Ser Lys Arg Val Leu Pro Asp Gly Arg Ile Glu Cys Tyr Ser Ser Ser 65 70 75 80 Thr Asp Gin Leu Glu Ala Pro Gly Thr Val Ser Glu Glu Ser Gln Val Leu Thr Asp Val Glu Ser Leu Ile Met Asp Asp Lys Ile Val Glu Asp $100 \ 105 \ 110$ Glu Val Asn Lys Glu Ser Val Pro Met Arg Glu Thr Val Ser Ile Arg 115 120 125 Lys Ile Gly Ser Lys Pro Arg Ser Ile Pro Pro Pro Gly Arg Gly Gln
130 135 140 Asp Tyr Arg Tyr Ser Gln Tyr Lys Arg Leu Arg Glu Glu Ile Asp Lys 165 170 175 Tyr Glu Gly Ser Leu Asp Ala Phe Ser Arg Gly Tyr Glu Lys Phe Gly 180 185 Phe Ser Arg Ser Glu Thr Gly Ile Thr Tyr Arg Glu Trp Ala Pro Gly 195 200 205 Ala Thr Trp Ala Ala Leu Ile Gly Asp Phe Asn Asn Trp Asn Pro Asn 210 220 Ala Asp Val Met Thr Gln Asn Glu Cys Gly Val Trp Glu Ile Phe Leu 225 230 240 Pro Asn Asn Ala Asp Gly Ser Pro Pro Ile Pro His Gly Ser Arg Val 245 250 255 Lys Ile Arg Met Asp Thr Pro Ser Gly Asn Lys Asp Ser Ile Pro Ala 260 266 270 Trp Ile Lys Phe Ser Val Gln Ala Pro Gly Glu Leu Pro Tyr Asn Gly 275 285 Ile Tyr Tyr Asp Pro Pro Glu Glu Glu Lys Tyr Val Phe Lys Asn Pro 290 295 Gln Pro Lys Arg Pro Lys Ser Leu Arg Ile Tyr Glu Ser His Val Gfy 305 310 315

Met Ser Ser Thr Glu Pro Val Ile Asn Thr Tyr Ala Asn Phe Arg Asp 325 330 335Asp Val Leu Pro Arg Ile Lys Lys Leu Gly Tyr Asn Ala Val Gln Leu 340 345 350 Met Ala Ile Gin Glu His Ser Tyr Tyr Ala Ser Phe Gly Tyr His Val 355 360 365 Thr Asn Phe Tyr Ala Ala Ser Ser Arg Phe Gly Thr Pro Asp Asp Leu 370 380 Lys Ser Leu Ile Asp Lys Ala His Glu Leu Gly Leu Leu Val Leu Met 385 390 395 400 Asp Ile Val His Ser His Ala Ser Thr Asn Thr Leu Asp Gly Leu Asn 405 410 415 Met Phe Asp Gly Thr Asp Gly His Tyr Phe His Ser Gly Pro Arg Gly 420 425 430 His His Trp Met Trp Asp Ser Arg Leu Phe Asn Tyr Gly Ser Trp Glu 435 Val Leu Arg Phe Leu Leu Ser Asn Ala Arg Trp Trp Leu Asp Glu Tyr 450 460 Lys Phe Asp Gly Phe Arg Phe Asp Gly Val Thr Ser Met Met Tyr Thr 465 470 480 His His Gly Leu Gln Val Asp Phe Thr Gly Asn Tyr Asn Glu Tyr Phe
485 490 495 Gly Tyr Ala Thr Asp Val Asp Ala Val Val Tyr Leu Met Leu Leu Asn 500 505 Asp Met Ile His Gly Leu Phe Pro Glu Ala Val Thr Ile Gly Glu Asp 515 520 525 Val Ser Gly Met Pro Thr Val Cys Ile Pro Val Glu Asp Gly Gly Val 530 540 Gly Phe Asp Tyr Arg Leu His Met Ala Val Ala Asp Lys Trp Val Glu 545 550 555 560 I ie Ile Gln Lys Arg Asp Glu Asp Trp Lys Met Gly Asp Ile Val His 565 575 Met Leu Thr Asn Arg Arg Trp Leu Glu Lys Cys Val Ser Tyr Ala Glu 580 585 590 Ser His Asp Gln Ala Leu Val Gly Asp Lys Thr Ile Ala Phe Trp Leu
595 600 605 Met Asp Lys Asp Met Tyr Asp Phe Met Ala Leu Asp Arg Pro Ser Thr 610 620

Pro Leu Ile Asp Arg Gly Val Ala Leu His Lys Met Ile Arg Leu Ile 625 $$ 630 $$ 630 $$ Thr Met Gly Leu Gly Gly Glu Gly Tyr Leu Asn Phe Met Gly Asn Glu 645 650 Phe Met Gly Asn Glu 655Phe Gly His Pro Glu Trp Ile Asp Phe Pro Arg Gly Asp Leu His Leu 660 665 670Pro Ser Gly Lys Phe Val Pro Gly Asn Asn Tyr Ser Tyr Asp Lys Cys 675 680 685 Arg Arg Arg Phe Asp Leu Gly Asn Ser Lys His Leu Arg Tyr His Gly $690 \hspace{0.5cm} 695 \hspace{0.5cm} 700$ Met Gln Glu Phe Asp Gln Ala Ile Gln His Leu Glu Glu Ala Tyr Gly 705 710 720 Phe Met Thr Ser Glu His Gln Tyr Ile Ser Arg Lys Asp Glu Arg Asp 725 730 735 Arg Ile Ile Val Phe Glu Arg Gly Asn Leu Val Phe Val Phe Asn Phe 740 745 750 His Trp Thr Ser Ser Tyr Ser Asp Tyr Arg Val Gly Cys Leu Lys Pro Gly Lys Tyr Lys Ile Val Leu Asp Ser Asp Asp Pro Leu Phe Gly Gly 770 775 780 Phe Gly Arg Leu Ser His Asp Ala Glu His Phe Ser Phe Glu Gly Trp 785 790 795 800 Tyr Asp Asn Arg Pro Arg Ser Phe Met Val Tyr Thr Pro Cys Arg Thr 805 815 Ala Val Val Tyr Ala Leu Val Glu Asp Glu Val Glu Asp Glu Leu Glu 825 830 Pro Val Ala Gly 835

- (2) INFORMATION FOR SEQ ID NO: 30:
 - (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2805 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

 - (1x) FEATURE: (A) NAME/KEY: CDS (B) LOCATION:131..2677
 - (x1) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

AGT	GAAT	TCG .	AGCT	CGGT	AC C	CGGG	GATC	C GA	ПСБ	CATT	TCT	CGCT	ATT	GCTT	TCCGTT	•	60
TAT	ПСС	ATA	TATA	AAAT/	AT C	AA ATI	CTAA	T CA	CTTG	CGCC	ATT	TCTA	TCT	стст	CCAAAC		120
TCT	CACCI	SAA (ATG (Met '	Val '	TAC Tyr 340	TAC A	ACT (Thr 1	GTA :	Ser (GGC A G1y 845	ATA (CGT Ang	TTT Phe	Pro (TGT Cys 850		169
GCA Ala	CCT Pro	TCA Ser	CTC Leu	TAC Tyr 855	AAA Lys	TCT Ser	CAG G1n	CTC Leu	ACC Thr 860	Ser	TTC Phe	CAT His	GGC G1y	GGT Gly 865	CGA Arg		217
AGG Arg	ACC Thr	TCT Ser	TCT Ser 870	GGC G1y	CTT Leu	TCC Ser	TTC Phe.	CTC Leu 875	TTG Leu	AAG Lys	AAG Lys	GAG G1u	CTG Leu 880	TTT Phe	CCT Pro		265
CGG Arg	AAG Lys	ATC Ile 885	TTT: Phe	GĊT Ala	GGA G1y	AAG Lys	TCC Ser 890	TCT Ser	TAT Tyr	GAA G1u	TCT Ser	GAC Asp 895	TCC Ser	TCA Ser	AAT Asn		313
TTA Leu	ACT Thr 900	GTC Val	TCT Ser	GCA Ala	TCT Ser	GAG G1u 905	AAG Lys	GTC Val	CTT Leu	GTT Val	CCT Pro 910	GAT Asp	GAT Asp	CAG G1n	ATT 11e		361
GAT Asp 915	GGC G1y	TCT Ser	TCT Ser	TCT Ser	TCA Ser 920	ACA Thr	TAT Tyr	CAA G1n	TTA Leu	GAA G1u 925	ACC Thr	ACT Thr	GGC G1y	ACA Thr	GTT Val 930		409
TTG Leu	GAG G1u	GAA G1u	TCC Ser	CAG G1n 935	GTT Val	CTT Leu	GGT Gly	GAT Asp	GCA Ala 940	GAG G1u	AGT Ser	CTT Leu	GTG Val	ATG Met 945	GAA Glu		457
GAT Asp	GAT Asp	AAG Lys	AAT Asn 950	GTT Val	GAG G1u	GAG G1u	GAT Asp	GAA G1u 955	GTA Va1	AAA Lys	AAA Lys	GAG G1u	TCG Ser 960	GTT Val	CCA Pro		505
TTG Leu	CAT His	GAG G1u 965	ACA Thr	ATT Ile	AGC Ser	ATT I le	GGA G1 y 970	AAA Lys	AGT Ser	GAA Glu	TCT Ser	AAA Lys 975	CCA Pro	AGG Arg	TCC Ser		553
ATT Ile	CCT Pro 980	CCA Pro	CCT Pro	GGC Gly	AGT Ser	666 61 y 985	CAG G1n	AGA Arg	ATA Ile	TAT Tyr	GAC Asp 990	ATA Ile	GAT Asp	CCA Pro	AGC Ser		601
TT6 Leu 995	AJ9	GGT Gly	TTC Phe	CGT Arg	CAG Gln 1000	His	CTT Leu	GAC Asp	TAC Tyr	CGA Arg 1009	Tyr	TCA Ser	CAG G1n	TAC Tyr	AAA Lys 1010		649
AGG Arg	CTG Leu	CGT Arg	GAG G1u	GAA Glu 1015	He	GAC Asp	AAG Lys	TAT Tyr	GAA Glu 1020	Gly	GGT Gly	TTG Leu	GAT Asp	GCA Ala 1029	Phe		697
ICT Ser	CGT Arg	Gly	TTT Phe 1030	Glu	Lys	Phe	Gly	Phe	Leu	CGC Arg	Ser	Glu	Thr	GGA Gly	Ile		745

ACT Thr	TAT Tyr	AGG Arg 104	Glu	TGG Trp	GCA Ala	CCT Pro	GGA G1y 1050	Ala	ACG Thr	TGG Trp	GCT Ala	GCA Ala 105	Leu	ATT Ile	GGA Gly	793
GAT Asp	TTC Phe 1060	Asn	AAT Asn	TGG Trp	AAT Asn_	CCT Pro 1065	Asn	GCA Ala	GAT Asp	GTC Val	ATG Met 1070	Thr	CGG Ang	AAT Asn	GAG G1u	 <u>8</u> 41
Phe 107	GGT Gly	GTC Val	TGG Trp	GAG G1u	ATT Ile 1080	Phe	TTG Leu	CCA Pro	AAT Asn	AAC Asn 108	Ala	GAT Asp	GGT Gly	TCA Ser	CCA Pro 1090	889
CCA Pro	ATT I le	CCT Pro	CAT His	GGT Gly 1095	Ser	CGA Arg	GTA Va 1	AAG Lys	ATA Ile 1100	Arg	ATG Met	GAT Asp	ACT Thr	CCA Pro 110	Ser	937
	ATC Ile			Ser					He					Gln		985
CCT Pro	GGT Gly	GAA Glu 1129	He	CCA Pro	TAC Tyr	AAT Asn	GCC Ala 1130	Ne	TAC Tyr	TAT Tyr	GAT Asp	CCA Pro 113	Pro	AAG Lys	GAG G1u	1033
GAG G1u	AAG Lys 1140	Tyr	GTG Val	TTC Phe	AAA Lys	CAT His 1145	Pro	CAG G1n	CCA Pro	AAG Lys	AGA Arg 1150	Pro	AAA Lys	TCA Ser	CTT Leu	1081
AGG Arg 1155	ATT Ile	TAT Tyr	GAA G1u	TCT Ser	CAT His 1160	Val	GGG G1y	ATG Met	AGT Ser	AGT Ser 1165	Het	GAG G1u	CCA Pro	ATA Ile	ATT Ile 1170	1129
AAC Asn	ACA Thr	TAT Tyr	GCC Ala	AAC Asn 1175	Phe	AGA Arg	GAT Asp	gat Asp	ATG Met 1180	Leu	CCT Pro	CGC Arg	ATC Ile	AAA Lys 118	Lys	1177
CTT Leu	GGC G1y	TAC Tyr	AAT Asn 1190	Ala	GTT Val	CAG G1n	ATC Ile	ATG Met 1199	Ala	ATT Ile	CAA G1n	GAG G1u	CAT His 1200	Ser	TAT Tyr	1225
TAT Tyr	GCT Ala	AGT Ser 1209	Phe	GGG G1 y	TAC Tyr	CAT His	GTC Val 1210	Thr	AAC Asn	TTT Phe	TTT Phe	GCA Ala 121!	Pro	AGC Ser	AGC Ser	1273
CGA Arg	777 Phe 1220	Gly	ACT Thr	CCT Pro	GAT Asp	GAT Asp 1225	Leu	AAG Lys	TCT Ser	TTA Leu	ATA Ile 1230	Asp	AAA Lys	GCT Ala	CAT His	1321
6AG G1u 1235	TTA Leu	GGG G1y	CTG Leu	CTT Leu	GTT Val 1240	Leu	ATG Met	GAT Asp	ATT Ile	GTT Val 1245	His	AGC Ser	CAT His	GCG Ala	TCA Ser 1250	1369
AAT Asn	AAT Asn	ACG Thr	TTG Leu	GAT Asp 125	Gly	CTG Leu	AAC Asn	Met	TTT Phe 1260	ASD	GGT G1y	ACG Thr	GAT Asp	AGT Ser 126	His	1417

	TAC Tyr	TTC Phe	CAC His	TCC Ser 127	GIY	TCA Ser	CGG Arg	GGT G1y	CAT His 127	His	TGG Trp	TTG Leu	TGG Trp	GAC Asp 128	Ser	CGC Ang	1465
	CTT Leu	TTC Phe	AAC Asn 128	Tyr	GGA Gly	AGC Ser	TGG Trp	GAG Glu 129	Val	CTA Leu	AGA Arg	TTT Phe	CTT Leu 129	Leu	TCA Ser	AAT Asn	1513
	SCA Ala	AGA Arg 130	Irp	TGG Trp	TTG Leu	GAA G1u	GAG Glu 1305	Туг	AGG Arg	TTT Phe	GAT Asp	GGT Gly 1310	Phe	AGA Arg	TTT Phe	GAT Asp	1561
1	3G6 31 y 131	Val	ACT Thr	TCC Ser	ATG Met	ATG Met 1320	Tyr	ACT Thr	CCC Pro	CAT His	6GG Gly 1325	Leu	CAG G1n	GTA Val	GCT Ala	TTT Phe 1330	1609
4	ACT Thr	GGC G1y	AAC Asn	TAC Tyr	AAT Asn 133	GAG Glu 5	TAC Tyr	TTT Phe	GGA Gly	TAT Tyr 1340	Ala	ACT Thr	GAT Asp	GTA Va 1	GAT Asp 134	Ala	1657
	STG /a1	ATT Ile	TAT Tyr	TTG Leu 1350	met	CTT Leu	GTG Va 1	AAT Aşn	GAT Asp 1359	Met	ATT Ile	CA C His	GGT Gly	CTT Leu 1360	Phe	CCT Pro	1705
(AG ilu	GCT Ala	GTT Val 1365	Thr	ATT Ile	GGT Gly	GAA Glu	GAT Asp 137(Val	AGC Ser	GGA Gly	AAG Lys	CCA Pro 1375	Thr	TIT	TGC Cys	1753
1	IT lle	CCA Pro 1380	Va I	GAA G1u	GAT Asp	GGT G1y	GGT Gly 1385	1 aV	GGA G1y	TTT Phe	GAT Asp	TAC Tyr 1390	Arg	CTC Leu	CAC His	ATG Met	1801
1	iCC 11a 1395	He	GCC A1a	GAT Asp	AAA Lys	TGG Trp 1400	He	GAG G1u	ATT Ile	CTT Leu	AAG Lys 1405	Lys	AGA Arg	GAT Asp	GAG G1u	GAC ASP 1410	1849
1	GG rp	AAA Lys	ATG Met	Gly	GAC Asp 1415	ATT Ile	GTG Va 1	CAT His	ACA Thr	CTC Leu 1420	Thr	AAC Asn	AGA Arg	Arg	TGG Trp 1425	Leu	1897
6	AA lu	AAA Lys	Cys	GTT Val 1430	Ala	TAT Tyr	GCT Ala	GAA G1u	AGT Ser 1435	His	GAC Asp	CAA G1n	Ala:	CTT Leu 1440	Va 1	GGT Gly	1945
G	AE Sp	Lys	ACT Thr 1445	He	GCA Ala	TTT Phe	TGG Trp	CTG Leu 1450	Met	GAC Asp	AAG Lys	GAC Asp	ATG Met 1455	Tyr	GAC Asp.	TTC Phe	1993
A	eτ	GCT Ala 1460	arg	GAC Asp	AGA Arg	CCA Pro	TCT Ser 1465	Ihr	CCT Pro	CTT Leu	He	GAT Asp 1470	Arg	GGA Gly	ATA Ile	GCA Ala	2041
Ļ	TG eu 475	His	AAA Lys	ATG Met	ATC Ile	AGG Arg 1480	Leu	ATT Ile	ACC Thr	Met	GGC G1y 1485	Leu	GGC G1y	GGA G1y	GAA G1u	GGA Gly 1490	2089

																-
TAT Tyr	TTG Leu	AAT Asn	TTT Phe	ATG Met 1499	Gly	AAT Asn	GAA G1u	TTT Phe	GGA Gly 1500	His	CCT Pro	GAG G1u	TGG Trp	ATT 11e 150	Asp	2137
TTT Phe	CCA Pro	AGA Arg	666 61y 1510	Asp	CGA Arg	CAT His	CTG Leu	CCC Pro 1515	Asn	GGT Gly	AAA Lys	GTA Val	ATT I le 152(Pro	GGG Gly	2185
AAC Asn	AAC ,Asn	CAC His 1525	Ser	TAT Tyr	GAT Asp	AAA Lys	TGC Cys 1530	Arg	CGT Arg	AGA Arg	TTT Phe	GAT Asp 1539	Leu	GGT G1y	gat Asp	2233
Ala	GAC ASP 1540	TAT Tyr)	CTA Leu	AGA Arg	TAT Tyr	CAT His 1545	Gly	ATG Met	CAA G1n	GAG Glu	TTT Phe 1550	Asp	CAG G1n	GCA Ala	ATG Net	2281
CAA G1n 1556	His	CTT Leu	GAA G1u	GAA G1u	GCC A1a 1560	Tyr	GGT Gły	TTC Phe	ATG Met	ACT Thr 1569	Ser	GAG G1u	CAC His	CAG G1n	TAT Tyr 1570	2329
ATA Ile	TCA Ser	CGG Arg	AAG Lys	GAT Asp 1575	Glu	GGA G1y	GAT Asp	CGG Arg	ATC 11e 1580	He	GTC Val	TTT Phe	GAG G1u	AGG Arg 1589	Gly	2377
AAC Asn	CTT Leu	GTT Val	TTT Phe 1590	Val	TTC Phe	AAC Asn	TTT Phe	CAT His 1595	Trp	ACT Thr	AAC Asn	AGC Ser	TAT Tyr 1600	Ser	GAT Asp	2425
TAC Tyr	CGA Arg	GTT Val 1605	61y	TGC Cys	TTC Phe	AAG Lys	TCA Ser 1610	Gly	AAG Lys	TAC Tyr	AAG Lys	ATT Ile 1615	Val	TTG Leu	GAC Asp	2473
TCG Ser	GAT Asp 1620	GAT Asp	GGC G1y	TTG Leu	TTT Phe	GGA Gly 1625	Gly	TTC Phe	AAC Asn	AGG Arg	CTT Leu 1630	Ser	CAT His	GAT Asp	GCC Ala	2521
646 61u 1635	His	TTC Phe	ACC Thr	TTT Phe	GAC Asp 1640	Gly	TGG Trp	TAT Tyr	gat Asp	AAC Asn 1645	Arg	CCT Pro	CGG Arg	TCC Ser	TTC Phe 1650	2569
ATG Met	GTA Va 1	tat Tyr	Ala	CCA Pro 1655	Ser	AGG Arg	ACA Thr	GCA Ala	GTG Val 1660	Val	TAT Tyr	GCT Ala	TTA Leu	GTA Va 1 1665	G1u	2617
GAT Asp	GAA G1u	GAG G1u	AAT Asn 1670	G1 u	GCA Ala	GAG G1u	AAT Asri	GAA Glu 1679	Val	GAA Glu	AGT Ser	GĀA Glu	GTG Val 1680	Lys	CCA Pro	2665
GCC Ala	TCC Ser	GGC Gly 1685	. *	GATA	AGAT/	ATT 1	ragt/	VAGAG	SG AT	rocco	TAA	A GC/	V GGA/	ATGG	•	2717
TTA	CCT	ITG (ATCT	GCAT	Π G/	ACG/	ACGT/	A TAT	TTGAG	ACT	GGA	NATC(AT /	ATGAC	CTAGTA	2777
GATO	стсі	AG /	GTCG	ACC	rg cz	VGGC/	ATG								•	2805

(2) INFORMATION FOR SEQ ID NO: 31:

- (1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 849 amino acids (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein (x1) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

Met Val Tyr Tyr Thr Val Ser Gly Ile Arg Phe Pro Cys Ala Pro Ser 1 15 Leu Tyr Lys Ser Gin Leu Thr Ser Phe His Gly Gly Arg Arg Thr Ser 20 25 30 Ser Gly Leu Ser Phe Leu Leu Lys Lys Glu Leu Phe Pro Arg Lys Ile $35 ext{ 40}$ Phe Ala Gly Lys Ser Ser Tyr Glu Ser Asp Ser Ser Asn Leu Thr Val Ser Ala Ser Glu Lys Val Leu Val Pro Asp Asp Gln Ile Asp Gly Ser 65 70 75 80 Ser Ser Ser Thr Tyr Gln Leu Glu Thr Thr Gly Thr Val Leu Glu Glu Glu 95 Ser Gln Val Leu Gly Asp Ala Glu Ser Leu Val Het Glu Asp Asp Lys 100 105 110 Asn Val Glu Glu Asp Glu Val Lys Lys Glu Ser Val Pro Leu His Glu 115 120 125 Thr Ile Ser Ile Gly Lys Ser Glu Ser Lys Pro Arg Ser Ile Pro Pro 130 135 140 Pro Gly Ser Gly Gln Arg Ile Tyr Asp Ile Asp Pro Ser Leu Ala Gly 145 150 160 Phe Arg Gln His Leu Asp Tyr Arg Tyr Ser Gln Tyr Lys Arg Leu Arg 165 170 175 Glu Glu Ile Asp Lys Tyr Glu Gly Gly Leu Asp Ala Phe Ser Arg Gly
. 180 185

Glu Trp Ala Pro Gly Ala Thr Trp Ala Ala Leu Ile Gly Asp Phe Asn 210 220

Phe Glu Lys Phe Gly Phe Leu Arg Ser Glu Thr Gly Ile Thr Tyr Arg 195 200 205

Asn Trp Asn Pro Asn Ala Asp Val Met Thr Arg Asn Glu Phe Gly Val 225 235 240

Trp Glu Ile Phe Leu Pro Asn Asn Ala Asp Gly Ser Pro Pro Ile Pro 245 250 255 Asp Ser Ile Pro Ala Trp Ile Lys Phe Ser Val Gln Ala Pro Gly Glu 275 280 285 Ile Pro Tyr Asn Ala Ile Tyr Tyr Asp Pro Pro Lys Glu Glu Lys Tyr 290 295 300 Val Phe Lys His Pro Gln Pro Lys Arg Pro Lys Ser Leu Arg Ile Tyr 305 310 320 Glu Ser His Val Gly Het Ser Ser Het Glu Pro 11e 11e Asn Thr Tyr 325 330 335 Ala Asn Phe Arg Asp Asp Met Leu Pro Arg Ile Lys Lys Leu Gly Tyr 345 350Asn Ala Val Gln Ile Het Ala Ile Gln Glu His Ser Tyr Tyr Ala Ser 355Phe Gly Tyr His Val Thr Asn Phe Phe Ala Pro Ser Ser Arg Phe Gly 370 380 Thr Pro Asp Asp Leu Lys Ser Leu Ile Asp Lys Ala His Glu Leu Gly 385 400 Leu Leu Val Leu Met Asp Ile Val His Ser His Ala Ser Asn Asn Thr $405 \hspace{0.25cm} 410 \hspace{0.25cm} 410 \hspace{0.25cm} 415$ Leu Asp Gly Leu Asn Met Phe Asp Gly Thr Asp Ser His Tyr Phe His
420
430 Ser Gly Ser Arg Gly His His Trp Leu Trp Asp Ser Arg Leu Phe Asn 445 Tyr Gly Ser Trp Glu Val Leu Arg Phe Leu Leu Ser Asn Ala Arg Trp 450 460 Trp Leu Glu Glu Tyr Arg Phe Asp Gly Phe Arg Phe Asp Gly Val Thr 465 470 480 Ser Net Het Tyr Thr Pro His Gly Leu Gln Val Ala Phe Thr Gly Asn 485 490 495 Tyr Asn Glu Tyr Phe Gly Tyr Ala Thr Asp Val Asp Ala Val Ile Tyr 500 505 Leu Met Leu Val Asn Asp Met Ile His Gly Leu Phe Pro Glu Ala Val 515 525 Thr Ile Gly Glu Asp Val Ser Gly Lys Pro Thr Phe Cys Ile Pro Val 530 540

Glu Asp Gly Gly Val Gly Phe Asp Tyr Arg Leu His Met Ala Ile Ala 545 550 560 Asp Lys Trp Ile Glu Ile Leu Lys Lys Arg Asp Glu Asp Trp Lys Met
575
575 Gly Asp Ile Val His The Leu Thr Asn Arg Arg Trp Leu Glu Lys Cys 580 585 Val Ala Tyr Ala Glu Ser His Asp Gln Ala Leu Val Gly Asp Lys Thr 595 605 Ile Ala Phe Trp Leu Met Asp Lys Asp Met Tyr Asp Phe Met Ala Arg 610 620 Asp Arg Pro Ser Thr Pro Leu Ile Asp Arg Gly Ile Ala Leu His Lys 625 635 640 Met Ile Arg Leu Ile Thr Met Gly Leu Gly Gly Glu Gly Tyr Leu Asn $645 \hspace{0.5cm} 650 \hspace{0.5cm} 650$ Phe Met Gly Asn Glu Phe Gly His Pro Glu Trp Ile Asp Phe Pro Arg 660Gly Asp Arg His Leu Pro Asn Gly Lys Val Ile Pro Gly Asn Asn His 675 680 685 Ser Tyr Asp Lys Cys Arg Arg Arg Phe Asp Leu Gly Asp Ala Asp Tyr 690
 700 Leu Arg Tyr His Gly Met Gln Glu Phe Asp Gln Ala Met Gln His Leu 705 710 720 Glu Glu Ala Tyr Gly Phe Het Thr Ser Glu His Gln Tyr Ile Ser Arg 725 730 735 Lys Asp Glu Gly Asp Arg Ile Ile Val Phe Glu Arg Gly Asn Leu Val 745 Phe Val Phe Asn Phe His Trp Thr Asn Ser Tyr Ser Asp Tyr Arg Val 765 Gly Cys Phe Lys Ser Gly Lys Tyr Lys Ile Val Leu Asp Ser Asp Asp 770 780 Gly Leu Phe Gly Gly Phe Asn Arg Leu Ser His Asp Ala Glu His Phe 785 795 800 Thr Phe Asp Gly Trp Tyr Asp Asn Arg Pro Arg Ser Phe Met Val Tyr 805 815 Ala Pro Ser Arg Thr Ala Val Val Tyr Ala Leu Val Glu Asp Glu Glu 825 830 Asn Glu Ala Glu Asn Glu Val Glu Ser Glu Val Lys Pro Ala Ser Gly

CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

- 1. A nucleic acid sequence encoding a polypeptide, the encoded polypeptide comprising at least a portion of the amino acid sequence shown in Figure 4 or Figure 13 which retains sufficient starch branching enzyme (SBE) activity when expressed in E. coli KV 832 to complement the branching enzyme mutation therein; at least 200 bp; and exhibiting at least 88% sequence identity with the corresponding region of the DNA sequence shown in Figures 4, 9, 10 or 13, operably linked in the sense or anti-sense orientation to a promoter operable in plants.
- A nucleic acid sequence according to claim 1, comprising nucleotides 21-2531 of the nucleic acid sequence shown in Figure 4, or a functionally equivalent nucleotide sequence which hybridises under stringent hybridisation conditions with the nucleic acid sequence shown in Figure 4.
- A nucleic acid sequence according to claim 1, comprising nucleotides 131-2677 of the nucleic acid sequence shown in Figure 13, or a functionally equivalent sequence which hybridises under stringent hybridisation conditions with the nucleic acid sequence shown in Figure 13.
 - A nucleic acid sequence according to any one of claims 1, 2 or 3 comprising a 5' and/or a 3' untranslated region.
 - A nucleic acid sequence according to any one of claims 1 o 4, encoding a
 polypeptide having the amino acid sequence NSKH at about residue 697.
- 6. A nucleic acid sequence comprising at least 200 bp and exhibiting at least 80 88% sequence identity with the corresponding region of the DNA sequence shown in Figures 4, 9, 10 or 13, operably linked in the sense or anti-sense orientation to a promoter operable in plants.
 - 7. A nucleic acid sequence according to claim 6, comprising at least 300-600bp.



- 8. A sequence according to claim 6 or 7, comprising a 5' and/or 3' untranslated region.
- A sequence according to claim 8, comprising nucleotides 688-1044 of the sequence shown in figure 9, and/or nucleotides 1507-1900 of the sequence shown in Figure 10.
- 10 10. A sequence according to claim 6, comprising the nucleotide sequence shown in Figure 10.
 - 11. A replicable nucleic acid construct comprising a nucleic acid sequence according to any one of the preceding daims.
- A polypeptide comprising at least a portion of the amino acid sequence shown in Figure 4 or Figure 13 which retains sufficient starch branching enzyme (SBE) activity when expressed in E. coli KV 832 to complement the branching enzyme mutation therein; encoded by a nucleic acid with at least 200 bp which exhibits at least 88% sequence identity with the corresponding region of the DNA sequence shown in Figures 4, 9, 10 or 13, operably linked in the sense or antisense orientation to a promoter operable in plants.
- 13. A polypeptide according to claim 12, in substantial isolation from other25 polypeptides.
 - 14. A polypeptide according to claim 12 or 13, having the amino acid sequence NSKH at about position 697.
- 30 15. A method of modifying starch in vitro, the method comprising treating starch to be modified under suitable conditions with an effective amount of a polypeptide according to any one of claims 12, 13 or 14.
 - 16. A method of altering a plant host cell, the method comprising introducing into the cell a nucleic acid sequence comprising at least 200 bp and exhibiting at least 88% sequence identity with the corresponding region of the DNA sequence shown in figures 4, 9, 10 or 13, operably linked in the sense or anti-sense orientation to a suitable promoter active in the host cell, and causing transcription of



5

the introduced nucleotide sequence, said transcript and/or the translation product
thereof being sufficient to interfere with the expression of a homologous gene
naturally present in the host cell, which homologous gene encodes a polypeptide
having SBE activity.

- 17. A method according to claim 16, wherein the host cell is from a cassava,
 10 banana, potato, pea, tomato, maize, wheat, barley, oat, sweet potato or rice plant.
- A method according to claim 16 or 17, comprising the introduction of one or more further nucleic acid sequences, operably linked in the sense or anti-sense orientation to a suitable promoter active in the host cell, and causing transcription of the one or more further nucleic acid sequences, said transcripts and/or translation products thereof being sufficient to interfere with the expression of homologous gene(s) present in the host cell.
- 19. A method according to claim 18, wherein the one or more further nucleic acid
 20 sequences interfere with the expression of a gene involved in starch biosynthesis.
 - 20. A method according to claim 18 or 19, wherein the further nucleic acid sequence comprises at least part of an SBE I gene.
- 25 21. A method according to claim 20, wherein the further nucleic acid sequence comprises at least part of the cassava SBE I gene.
 - 22. A method according to any one of claims 16-21, wherein the host cell is selected from one of the following: cassava, banana, potato, pea, tomato, maize, wheat, barley, oat, sweet potato or rice.
 - 23. A method according to any one of claims 16-22, wherein the altered host cell gives rise to starch having different properties compared to starch from an unaltered cell.
 - 24. A method according to any one of claims 16-23, further comprising the step of growing the altered host cell into a plant or plantiet.

35



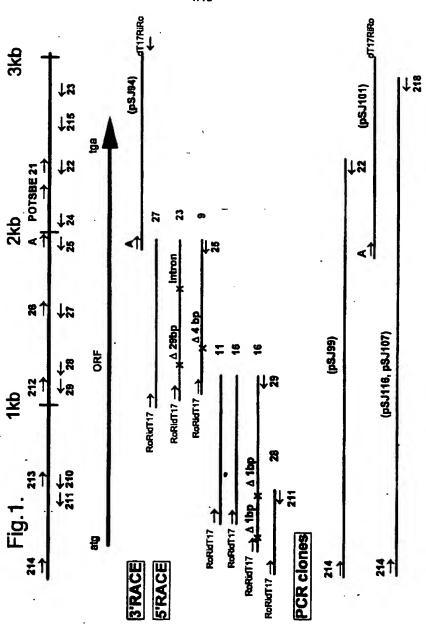
- 25. A method of obtaining starch having altered properties, comprising growing a plant from an altered host cell according to the method of claim 24, and extracting the starch therefrom.
- 26. A plant or plant cell into which has been artificially introduced a nucleic acid sequence comprising at least 200bp and exhibiting at least 88% sequence
 identity with the corresponding region of the DNA sequence shown in Figures 4, 9, 10, or 13, operably linked in the sense or anti-sense orientation to a promoter operable in plants, or the progeny thereof.
- 27. A plant according to claim 24, altered by the method of any one of claims 15 16-22.
 - 28. Starch obtainable from an extered plant according to claim 26 or 27, having altered properties compared to starch extracted from an equivalent but unaltered plant.
 - 29. Starch obtained from an altered plant according to claim 26 or 27, having altered properties compared to starch extracted from an equivalent but unaltered plant.
- 25 30. Starch according to claim 28 or 29 obtained from an altered plant selected from the group consisting of: cassava, banana, potato, pea, tomato, maize, wheat, barley, oat, sweet potato and rice plants.
- 31. Starch according to any one of claims 28, 29 or 30, having increased
 30 amylose content compared to starch extracted from an equivalent but unaltered plant.

Dated this 21st day of December 2000

NATIONAL STARCH AND CHEMICAL INVESTMENT HOLDING CORPORATION By their Patent Attorneys COLLISON & CO







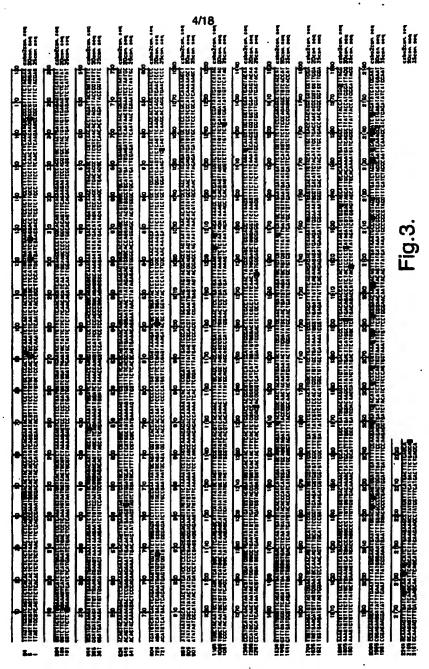
SUBSTITUTE SHEET (RULE 26)

Fig.2.

ры ' ' Э' '
ATBRATICACATCRATACTACTACTCACTATACCCATTTCTTTTTTTT
TACCTANCTETARCTATTATCCTENSTCATATCCCTANAGANANAMANAMANAMANATTTTTTTCANCTTETACCTTAATCANCCCAGTCANGAETETRARAGACTTGANGAC
HOLVASTLTLSLTS
Not
SCRAMATOCOMENCIAL RECENTATION CONTINUES TO THE TOTAL TOTAL TOTAL TOTAL CONTINUES TO THE PROPERTY OF THE PROPER
COCTITACCE TOTUSTO TOTAL TRANSPORTATOC SAMAGE LACACOLOGICA COLOGICA SOCIETY MANGET MADE TARACT ACCITACT TOC TOTAL TRANSPORTATION OF THE ACCITACT ACCITACT TOCAL TRANSPORTATION OF THE ACCITA
RESTLUMITALES TPOLBITELPY RENTTISE I REPELAPLACIOS TERREDRATES CLAPHE
MEAGEBEERTTTCTAGEAGERTCTTCTCTGBAAAGTCATCTCATEAAICTBACTCCTCAAAATGTAATGETCACTGCETCTBAAAAAAGTCCTTCCTBATGBTCBBATGCATT
TETTECHECHEANAMENTETECHEGAMBAGACCTITEAGTAGACTACACTAGACTGAGGAGTTTACCACTACCTAGACGAGTTTTETEAGGAAGGACTTACCACTTACCACT
TO THE TEACHT AND THE TEACHT AND TO THE TEACHT AND A THE TEACHT AND THE TEACHT AND A TO THE TEACHT AND A TEAC
THE PROPERTY AND THE PROPERTY AND THE PROPERTY AND
B B B T O O L E A P G T V S E E B Q V L T D V E S L I N D D E I V E D C V M S E
pani Pani
: The technical process in a catalog and an all technical and experient expersions and an analysis at a capacity feet
MCAMESTACCECTESTEATCATCCTACCCTTTTTAACCTACATTTCCTTCCACTAMONACTACATTTCCTTCCCTTTCTTATATACTETATCTACTTCCAAACCA
5. T P R R T T S I C K I C S K P R S I P P P G R G R B T B D P R L T G P R
dell.
ACACCTABATTACESSTATYCACANTACAAAAGACTCCGAGAAGAAATTGACAAGTATGAAGTATGAGTGCTGGATGCATTTTCCTCGCGCAGTGAA
TETBEATCTAAT SEECATAAGTETEATETTTTE TEAGES TETTCTTTAACTETTEATACTTEEATEAGAGETAAGTAAGAGCACCEATACTTTTEAAAGCAAAGAGTGCGTCACTT
) H L D V N V S O V K R L R E E I O K Y E G S L D A F S R G V g'K F G F S R S E
Bell
TY T
AGGANTANETTATANAGENTONOGENEGANGANTEN METGROCETECHTE ATTANAGENTE ATTANAGEN METANAGEN METANAGEN ATTANAGEN METANAGEN ME
AGGNATIAN TITATAGNAS TOGGCACCIAGAGC INCAS SECCITICAN TEANTOGRAF TITCAATAGCTGGAATCC TIATGCAGATGTGATACTCAGAATAGCTGTEST TITCGGGAG
TECTTATTEANTATETETEACEEGTBB1EETEBATECACECCACCTAACTAACETETAAABTTATTEACETTABBATTACBTCTACBACHAETEACACACEGTC
TECTTATICATION TO THE CONTROL OF THE PROPERTY OF THE PROPERTY OF THE CONTROL OF T
TECTIATICATIVE TECHNICOS TRETECISTE CALCULAR CARETIC TAMBITATICATE THE CONTROL TRETECIS TRANSPORTE THAT CALCULAR ACCACUANCE TO CONTROL TO THE CONTROL OF THE
TECTTATIONAL ATTENDED TO THE COLOR CONTROL CON
G I T Y N E W A P G A T W A C L L G O P N N W N P N A D V N T O N E C B W V C Name 100
TECTTATIONAL ATTENDED TO THE COLOR CONTROL CON
G I T Y N E W A P G A T W A C L L G O P N N W N P N A D V N T O N E C B W V C Name 100
COLITY BE WAP CAT WALL GOPHHWHIP HAS WHIT OHE CEVY C ADDITIONS COLITY BE WAP CAT WALL GOPHHWHIP HAS WHIT OHE CEVY C ADDITIONS COLITITIES COMMANDATE CONTRACT COMMANDATICS CANDEST CONTRACT COMMANDATICS CONTRACT COMMANDATICS CANDEST CONTRACT COMMANDATICS CONTRACT
G I T Y R E W A P G A T W A C L L G O P R R W R P R A B W R T O R E C B W W C AND DOG I CTITITIESSEANTANISCASATESSTEASCEACECASTANCIASSTEAST CONTROL CARACTER TRANSPORTATION CONTROL
TECHTATERATATETETACECTESTECTECSTECTECSTECACCACCTARTACTACCTECTARSTATTERACTTAGGATACCTTCCACCACCACCACCACCACCACCACCACCACCA
TECHTATERATATETETACCOTORIECT CERTECACCCACTATACTACCT CARACTATETECCT TARGET TATELACCT TARGET TATELACCT CARACTACT TATELACCACT CARACTACT CAR
G I T V R E V A P G A T V A L L G G P R H V R P H A D V H T O H E C B V V C ARROL JOD I CITITITECCANTANISCACACTACCACCACCACCACCACCACCACCACCACCACCA
G I T V R E W A P G A T W A L L G G P R R W R P R A D V R T O R E C E W V C AREA JOST TO R E W A P G A T W A L L G G P R R W R P R A D V R T O R E C E W V C AREA JOST TOTAL CONTROL OF THE PROPERTY OF THE
SEMANATORE CONTRETATA CONTRETA
S I T V R E W A P G A T W A L L G G P R R W R P R A D V R T O R E C R W V C AND 1 JON 1 CTTTTTRECCAATAATGECATCACACCACTACTACCACCACTACTACACCACCACTACT
TECHTATERATATECTEACECOTESTECTECRATECACCCACTATACTACCTCCACTATACTACCCTTATACTAC
S I T V R E W A P G A T W A L L G G P R R W R P R A D V R T O R E C R W V C AND 1 JON 1 CTTTTTRECCAATAATGECATCACACCACTACTACCACCACTACTACACCACCACTACT
TECHTATERATATECTEACECOTESTECTECRATECACCCACTATACTACCTCCACTATACTACCCTTATACTAC
G I T V R E V A P G A T V A A L L G O P R H V R P R A D V R T O H E C B V V C ARROL JOD I CITITITECCANTANISCACATGRITCACCACCANTICCCCAN GRITCIC CAMBRICANAGATACCA TERNATAR CONTINUANA CONTI
TETATEMENTAL ACCOUNTS TO THE CONTROL OF THE CONTROL
PRO 1 T Y R E W A P G A T W A L L G O P R H W H P H A D V H T O H E C B V V C ARROL JOD 1 CTITITECECANTANDSCACACTECTCCCCCASTCCCCCCASTCCCCCCASTCCCCCCCC
PROD JON 1 TO A T W B E W A P G A T W A A L L G O P H H W H P H A D V H T O H E C B V V E ARROL JON 1 CONTINUED ANTAINCAGATGROTICACCAC ANTICCCCAL RESTORTICATE ANAGATACTAC ATTERATACT TO THE ANAGATACT TO THE ANAGATACT ACCORDINATE ANAGATACT ANAGATACT ANAGATACT ANAGATACT ACCORDINATE
G I T Y R E W A P G A T W A A L L G O P R N W N P R A D V N T O N E C 8 V V PRINCETATION PRODUCTION OF COMPANY
TETATATIATE TETATATA A CALIGO PHINT HORIZOTA A TARGET CALACTAR TARGET CALACTAR AND THE CONTROL AND THE CONTROL AND A CALIGO PHINT HORIZOTA AND

Fig.2 (Cont).

9 (00).	
ATOTACACCEATEATECACTICACTICACTICACTICACTACAATCAATCATCTITGCATATCCACCTGATCTCTCTCTCTCTCTCTCTCTCTCT	,
TACATETECETACTACCTACCATCATCATCTACCTCCATCTACTACTACT	1680
NTTHELOVOFTENTNETFETATBVOAVVILALEMONIME	
CTETTICECARAGECTOTEACCATTESTGAGAGATGTTAGTGGAATGCCAACACTTTECATTCCGGTTQAGATGGTGGTGTTCCTTTAGTTATCGTGTCCACATGGCTGTTCCTQATAA	
BARAGGGTEYECGACABTGGTAACCACTTCTACAATCACCTTACGETTGTCAAAGGTAAGGT	1800
L F P E A V J I S E O V S G M P T V C L P V E D G G V G F D V R L M N A V A D C	
· · · · · ·	
Action 1	
TREETTGAGATTATYCAGAAGAGAGATGAGATTGGAAAATGGGTGACATTGTACATATCCTGACCAACAGCCBGTGGTAGATTGTTGTTATCCTGAAAGTCATGACCAGGC	
ACCEMACTETAMTANGTETICTETETICTACTTCTAACCTTTTACCCACTGTAACATGTATACGACTGGTTCCCCCCACCACCTTTTCACCAAAAAAATACCACTTTCACTACT	
VYETIOS PPED WENGD) V H PLT BER VLEEC VETA E EM A O A	
CITETTES GACAAAAC TATIGCATITIGGE IGATUGACAAGGALAIGIA IGAC I KEATGETECGYGACAEACCATC TACTEC TE LITATAGATCGTGGAATAGCATIGCAC AAAATGAT	2040
GAREAGE DETTTENTAL TARGETANAL CONCENTRATE CONTINUE TO A C	2040
L F G D E T 1 A F W L M D K O M Y D F N A R D E P 3 T P L 1 D R C 1 A L M E N 1	
Non t	
1	
AGGETTATTATCATGGCCTTACCCCOMGAAGCATATTTBAATTTTATCGGAAATBAATTTGGACATCTGAGGGGATGGATTTTCCAAGAGGMGATCGACATCTGCCCAATBSTAAABTA	2180
TECEPATA TOUR PROPERTY SECOND TO THE SAME AND THE SAME PARTY OF THE SAME PARTY SAME PART	
RLITH GLGGEOVLYFN GXEFOHPEW I BFPRED ONLPWERY	
prom v	
ATTECAMBAACAACCACAGTTATGATAAATGCGGTGGTACATTTGATCTAGGTGATGCAACTATCTAAGATATCATGAAATGCAAGGGTTTGATCAGGCAATGCAACATCTTGAACA	2280
TANGGTCCCTTGGTCCAATACTATTTACGCCACCATCTAAACTACATCCACTACGTCTATACAACTACGTTACGTTCCGTTCCGAAACTACTACTCCGTTCCGTTACGTTC	
ІР СИМ И Б У ОКС ПРЕГО L СО А ОТ L П У И С И О Е ГО О А И О И L E E	
ACCTATOSTTTCATGACTTCTEACCACCAGTATATATCACCACACCATGAACGAGATCGDATCATTGTCTTTGACACCCCAACCCTGTTTTTTTCACCTTTCATTGACTAACAGC	2900
CHIATACEARACTACTGARACTCGTBETCATATATAGTGCCTTCTACTTCCTCTAGTGCCTAGTAACAGCACCTCTTTTGCAACAAAAACATAACTTGAAACTAACT	
TATTCAGATTACCEAGTTEGGTRECTTCAAGTCACCAAAGTACAAGATTETTTTGGACTCCGATGATGGCTTGTTTBGACGCTTCAACAGGCTTAGTCATGATGCCGAGCACTTCACCTTT	
	2620
ATAABTETAATGGETCAACCGACGAAGTTEAGTCCTTTEATGTTCTAACAAAACETCACCTAGTAGGGAACAAGCTGGGAAGTTGTGGBAATGAGTACTACGGCTGGTBAAGTGGAAA	
V B B V R V G C P E S & X V X I V L O S D O G L P G G F M R L S M D A E M P T F	
CALEERTGGTATGATAACCGGCCTEGGT(CTTCATGGTATATGCACCATCTAGGACAGCAGTGGTCCATGCTTTAGTAGAAGATGAAGAAGAAGAAAGA	
CTOCCCACCATACTATTCCCCGGACCCAGGAAGTACCATATACGTCCTAGATCCTGTCGTCACCAGGTAGAAATCATCTTCTACTTCTTCTTACTTCCTCTTACTTCATCTTCACTTTCACTT	5040
DAYEM RPRSEN TAPES BEAT A VERY BALVE BEFERE BEAT WERE	
Baraki Masa B -	
GTBAAACEAGCTCCGGCTBAGATAGATATTTAGTAAGAGGATCCCCTAAAGCAGGAATGRTTAACCTGTGCATTGAACGACGTATATTGAGACTTGAAATTGATTTGCTGCTCA	2760
cacyttertergageceactetatetataaatcaytetectageebatticetectaecaatt <mark>eracetagacetagetagetacetectecatetaacttaact</mark>	
Y E P A S G	
••	
Sent mile such	
June 1 hay 1 hay 1	
pad i pad i Great adapt attaattelaagetelageerrargaraticatatelageerrargaratitetaaatelttäraatelttäreareteerretetaaattatat	
Pad I GRACHERARDATATTANTTECANGGETEANGGERDAGATACMGGCCATANTGCATGATGATATATAGAAGCTECECANGTTETANATEANTTAREAMBETBEGTTGCAGTETGTAMATTATATA BETBITGTTATAMTTANGGTTCCGAGTTCCTATGTGCCGGTATTACGTACTACTATACTTTCGABGGGTTGAMCATTTAGTALATEGTTCAMAGGAGGTTAGAAGATTTAATATACT	
ectbitictatatataattaaeetteebaetteegtetetateteegetattaeetaetaetataettitegabbogttbaacatttagtaaa tegetebaebeaebtba	. 2890
	. 2030
ectbitictatatataattaaeetteebaetteegtetetateteegetattaeetaetaetataettitegabbogttbaacatttagtaaa tegetebaebeaebtba	. 2000
CCTRIBICITATAATTAAGETICCGAETICCGGTTTTACGTACTACCTACTACTACTACTTCGARGGGTGAACATTTAGTGAATGGTCGACGGCGGTGAGACATTTAATATAC Books	2000
ectbitictataattaaesticegaeticegtetetatoloogtattacetactaltaluctticumbogtugaacatttagtalaetegetegaegteagaeatttaata Boo 1 Tabiactticoeagueacgivaltatoggataccatoggatgiceoctaggaaaaaittigibiatacoectactagaettitaaatetegeatotteeacataaagtogtogtugaato	2030
ectbitictataattaaesticegaeticegtetetatoloogtattacetactaltaluctticumbogtugaacatttagtalaetegetegaegteagaeatttaata Boo 1 Tabiactticoeagueacgivaltatoggataccatoggatgiceoctaggaaaaaittigibiatacoectactagaettitaaatetegeatotteeacataaagtogtogtugaato	2630
PETBTTGTTATATTAAGTACEGAGTTCCGAGTTCCTATGTCCGGTATTACCTACTACTATATATCTTCGAGGGGTGAACATTTAGTAAATCGGTCAGAGCACGTGAGACATTAATATAC POB 1 ***********************************	3000



SUBSTITUTE SHEET (RULE 25)

Fig.4.	pitco I
CTETETARETTERAGGAAATEGERCACTACACCATAICAGGAATACGTTTICCTTGTGET	
	P L C X 8 0 S T G F H G Y N - R T & B C
TTTCCTTCAACTTCAABCACCCGTTTCTAGGAGGGTCTTCTCTGGAAACTCATCTCATGA	TCTGAETCCTEARATGTAATGETEAGTGETTCTAAAAGAGTCCTTEGTGATGETCERA
AAAGGAAGTTBAAGTTCCTCCCAAAAGATCCTCCCAGAAGAGACATCTTCAAGTAGAGTACTT	
	5 0 5 9 H F M V T A B K R V L P O G R
TTERATECTATYCTTCTTCACACATATCATTGGAAGCCCCTGCCACACTTTCAGAACATCC AACTTACGATAACAACAACTTGTCTAGTTAACCTTCCCCCACCCTGTCAAAGTCTTCTTAG	
	0 * 1 D * E S 1 M 0 D 4 1 * E D E
Time 1) tired 00
TAMATAMABANTETETTECAN TECCOCABACAGITACCAT CAGAAAAATTEGATETAMECA ATTTATTTETTAGACAAGGITACOCCCTETGTEAATCETAGTETTTTTAACCTAGATTTGGT	
	R S 1 P P P G R G G R 1 T O L D P S L
Note in:	ושק
CAGECTTTCGTCAMCACCTARATTACCEGGTATTCACAGTACAAAAAACTCCCGAGAAGAAATT	
T G F B G H L D Y R Y S G Y L R L R E E 1	0 K 7 E 6 S L G A F S R G Y & E F 6 F
CACOCASTRAACASSAATAACTTATASASASTOSSCACCASSASCTACGTCSCCTUCATTS	ATTORNBATTTEAATAMETORAATEETAATGEAGATOTEATGACTEAGAATBAGTBIG
STRUCTURE TO THE PROPERTY OF T	TAREETE JAART STEEDEN THE STEEDEN THE SOUTH STATE TO THE STEEDE STATE ST
8 R 8 C T G 1 T Y R E V A P C A T V A A L	1 & B F # H W M P M A D Y M T O H E C
	1 6 9 7 11 1 1 11 11 11 11 11 11 11 11 11 11
	1 6 9 F Я И Ф И Р Я А О У Я Т О И Е С 201
GTBTETBERRATETYTTTEECEAATAA ISEAGATGBTTEACACCACCAA I TECEEATGETTET	CONTINUED TATES TAREAL SERVICE STREET SERVICES S
GTBTCTBCDABATCTTTTTGCCCAATAATGCAGATGBTTCACCACCACTACTTCCCCATGGTTCTT CACAGACCCTCTAGAAAAACGGCTTATTACGTCTACCAAGTGGTGGTTAAGGGGTACCAAGA	COMTAMORATACCCATGORATACTCCATCTOCCACAAABATTCTATTCCTGCTTGBA MO EXTERNITIETATGCCTATGMODIABACCCTTGTTTCTAAGATAAGCCGAACCT
GTRICTECCARATETITITECCAATAATECAGATERTICACCACCACTATICCCCATENTETT CACAGACCCTCTAGAAAAACCCCTTATTACGTCTACCAAGTGGTGGTTAACCGGTACCAAGA G V V E 1 F L P H H A O G S P P I P H Q S	COMPLIANCE TATES CONTROL TO THE CONTROL TO THE CONTROL OF T
GTBTCTBCDABATCTTTTTGCCCAATAATGCAGATGBTTCACCACCACTACTTCCCCATGGTTCTT CACAGACCCTCTAGAAAAACGGCTTATTACGTCTACCAAGTGGTGGTTAAGGGGTACCAAGA	COMPANAMENTACCATIGNATACTICATIC TROCAMANAMENTIC TATTIC CONTROL ON DESCRIPTION OF THE STATE OF THE
GTRICTEGRALATEUTTITECCEATANICE DATERTICACCACETATICCCCATEGITET CACMACCCTCTARAMANCOSCTTATTACETCTACCAGETGRIGGTTAACEGGTACCAMAN C V V E 1 F L P U N A O G S P P 1 P N Q S TEAMSTTETEARTICAAGCACCAGGITEAACTCCCATATAATGGCATATACTATACTCTCCC	CONTINUENCE TO THE PROPERTY OF
GTRICTEGRALATETTTTGCCGARTAITECGATERTTCACCACCACTATTCCCCATERTTCT CACAGACCCTCTAGAAAACGCCTTATTACGTCTACCAACTGGTGGTGGTAACAGCGTACCAAAACGCCTCCCAAACGGTACCAACACTGGTGGTAACAGCGTACCAAAACGCCCCAACAACGGTGAACCACCAACAACGCCCAACAACACCCCCAACACACAC	CONTINUENCE TO THE PROPERTY OF
GTRICTEGRALATETTTTGCCGARTAITECGATERTTCACCACCACTATTCCCCATERTTCT CACAGACCCTCTAGAAAACGCCTTATTACGTCTACCAACTGGTGGTGGTAACAGCGTACCAAAACGCCTCCCAAACGGTACCAACACTGGTGGTAACAGCGTACCAAAACGCCCCAACAACGGTGAACCACCAACAACGCCCAACAACACCCCCAACACACAC	CONTINUENT ACCORDINATION TO THE CONTINUENT ACCORDINATION TO ACCOR
GTRICTEGRALATETTTIGECEARIATICECATETTETCACACEGATICACCACETETTETCACACEGATICACCACETATICCCCATESTETCACCACEGATICACCACEGATICACCACEGATICACCACEGATICACCACEGATICACCACEGATICACCACCACEGATICACCACCACCACCACCACCACCACCACCACCACCACCAC	CONTINUENT ACCORDINATION TO THE CONTINUENT ACCORDINATION OF TH
GTBTETEGRAMATECUTTITICECEANIANTEERDATERTICACCARATICECCATEGITET CACMMACCETETARAMANCOSETTATTACETETACAMETERICUTTAAGEGSTACCAMAN C V W E 1 F L P W M A O G S P P 1 P M G S TEAMSTTETEARTICAAGCACCAGSTEAACTCCCATATAATGCCATATATATATCCTCCC AGTTEAAGATCAGATCAGTTCGTESTCCACTTAAGGGTATATTACCATATATATATATCCTCCC GTTATAGGTCGCCACGTTGGAATGAGTAGGTACGTACCTAATTAACACATATCCCAACTTT.	CONTINUENT ACCUMENTATION OF PART OF PA
CTRICTEDIDATE TYPETECCAATA FECHA FER THE ACCACE A FICCECA TENTET. CAMMACCET TARBAMAACOSE TRATTACTET ACCAME TORTOUT FAREGOS MACHAMATOR C V V E 1 F L P N N A O G S P P 1 P N G S TECAMETTET CAMECACE ACCUST CONTRACTOR OF TARBAT ACCOUNT A TORTOUT CONTRACTOR OF THE ACCOUNT A TORTOUT A	CONTINUENCE TO PROPERTY OF THE TARGET TO TARGET TO THE TARGET TO T
CHRICTEDDALATETTTTECCAATATECTACTATETTTACCACCAATTCCCCATESTETT CACAGACCCTCTAGAAAACCCCTTATTACGTCTACCAAGTGGTGGTGATACCACCAATTCCCCATEAAAACCCCTTATTACGTCTACCAAGTGGTGGTGATACTACTACCACCAATAACCCCTATATACACCTCCCCCACTATAAACACCTCCCCCC	CARTIMADATACCATIGNATACCCATTERCAACAARATECTATTCCTGETTERA EXTERT SCHAFFER CARTIMINIA CONTINUE CO
CTRICTEDIDATE TYPETECCAATA FECHA FER THE ACCACE A FICCECA TENTET. CAMMACCET TARBAMAACOSE TRATTACTET ACCAME TORTOUT FAREGOS MACHAMATOR C V V E 1 F L P N N A O G S P P 1 P N G S TECAMETTET CAMECACE ACCUST CONTRACTOR OF TARBAT ACCOUNT A TORTOUT CONTRACTOR OF THE ACCOUNT A TORTOUT A	CARTIMADATACCATIGNATACCCATTERCAACAARATECTATTCCTGETTERA EXTERT SCHAFFER CARTIMINIA CONTINUE CO
CTRICTEDDALATETYTTECCCAATATCECEAGTESTICACCACCTATICCCCATCSTICTT CACAGACCCTCTAGAAAACCCCTTATTACGTCTACCAAGTGGTGGTGATACCCCCATCAAAACCCCTTATTACGTCTACCAAGTGGTGGTGATACTACCACCAAAACCCCCAATACACCCCCCCC	CONTINUADA ACCUATGRA TACCOLATERACIA ANABATTE TATTOCTGETTERA BY KIR R R D T P & 8 W K B S I P A W CASSAGRAMMAN TARGET TEANANTET TRACCCAAAAAN CAARACCAAAA TCACTTCORA SED CICCITCA TACACCAARTITI TACACCAAAAAACCATCACTACTACAAA CICCITCA TACACCAARTITI TACACCAAAAAACCATCACTACTACTACTACTACTACTAC
CHRICTEDDALATETTTTECCAATATECTACTATETTTACCACCAATTCCCCATESTETT CACAGACCCTCTAGAAAACCCCTTATTACGTCTACCAAGTGGTGGTGATACCACCAATTCCCCATEAAAACCCCTTATTACGTCTACCAAGTGGTGGTGATACTACTACCACCAATAACCCCTATATACACCTCCCCCACTATAAACACCTCCCCCC	THE

R/15

Fig.4 (Cont).

4	T	GA.	ΙŒΙ	*	84	CTC	TÇD	CT.	111		CTAT	COL	ac.	TOC	SAE	211	ETA	AGE	1111	m	777	AAA	OÇA		TRET	-	TTCQ.	ATE.	MI	CA	STT:	TEAT	1821	m	104	771	BAT	LEGS	
7	AF	-	~	č	F	gar.	400	5	444	177	DAT/	~	TEA	N.C	ett		EAT	120	AAN			111	cer	TCC	1004	~	MEE	IAC	TEA		244	ei.	~	-	=		***		1440
							_			•••					•••	•	•	•••	,												-				•••				
*	٧	•		v	0			L		N	¥	6		٧	E	•	L	1	•	L	L					¥	. 1	•				Ö	Ġ	•		F	0		
		•	-44	TO:		772	-	٠,	TEAT	ree	ATTC	***	-014	211	777	we	ccc		uei	MT	LAAT	etti	-	TAT	BEAR	e Ti	LATE			TET	207	TAT	776	MTE	***	***	. AAT	RATA	
_	_	_		-		_	-	-	-	_	-	_		-		-	-	_		_	_	-	-	_	_	-	_	_	-	_	-	-		_	-		-	-	1980
A	TE	•	111	æ	TA	CAT	TO	31	ACT	LCE	TAAC	ere.	CAT	214	***	TEG	CCE	TTE.	ATET	ITM	TVA	-	CET	ATA	COTT	w	TAG	H		ME.	CCA	MTA		TAC	ш	AAC	TTA	ETAT	
_	_			_		_	_			_		_	_	_			_	_	_		E 1		_	_							-	_		_			_	_	
۰	T		•	•		•	•	*	•	•		0	٠	,	•	•	•	•	•	•	. '	•	٠	•	•	•	•	' '	•		•	•	٠	•	٠	•	•	•	
		_		_																				***				-		700			***		_	- 20		DE TO	
_	_	÷	•	•	_		_	_	_	_	-		_	-	_		-+	_	_	_		-	-	_		-	~-	_	_		_	_	_	-	_	_	_	_	1688
ж	TA			Ç.M		~		cr	cce	E.	2782	TAR	CC.	CIT	CTA	EAA	TEA	ter	TACE	101	810	***	TAA	080	CTTC.	111	TAC	MO		MED		ICT/	411	ec.	246	216	TAC	CEAE	
														_			_	_		_				_		_					_	_	•	_			_		
	•	•	•		L	•	•	E	4	٧	T	٠	4	E	•	7		Ç		P	т,		1	,	•	E	D 1	•			•		Ŧ		L				
														.																									
_			_	•	_	-				_	_	_		_	_	_	-	-	-	_	1761	_	_	_		•	_	_	-	_	-	_	_	_	_	_	_	_	1600
M	ce	5	TAT	11/	ğ	CA	KI	TA	414	Q T	2770	TCT	CTA	271	LT2	ACC.	111	I AC	CCAC	TO	ME	TET	TAC	erc	TEST	TEI	CCCI	CM	Ċ	CET	m	بغن	CAI	424	ATĀ	CGI	ETT	8431	
٠		1	•	E	¥		£	ı		٥	E		0	C	٥	¥	ĸ	H	•	Ð		- 11		L				١,	,		ĸ	¢	۳	8	٧	A	£	8	
AT		×	-	ccs	CTI	TST	1221	CA		uc	BAFT	CÇ A	177	TEE	212	ATO	GAC.	LAC	CAT.	1781	1101	erre	ATE	ec 1	377	AC!	GAC	ATI	TAE	TCC	ren:	LTA:	GAT	CET	CEA	OTA	OCA'	TEE	
Ξ	=	Ξ	Ξ	Ξ		-	=	_		_	***					740	-	777	741		TACI	24.24	***			7		72	<u></u>	Mr	S L	747	774	~	77	PAT	***	<u></u>	1530
		-	ш	***	_	-	u.c.		.,,,	•					-	-				-					_	•••							• • •		•				
	_									7						_										•												١,	
•	•	•	•	-	٠	•	•		•	•	•	•	•	•	٠	•	•	•	•	"	٠.	•	-	-	•	•		•		•	•	•	•	-	•	٠	-	•	
			_	.,					-																														
			г	••					_	•			•																										
м	*	41	res.	164		ET	AT	40	CATI	ŧ.	ATTA	440	204	BAA	GÇA	TAT	TTP	MT	1111	LTO	200	سة	m	CEM	CACC	ccı	LASTI	1041	πu	m	TCE/	w	CET	BAT	CTĀ	LAT	2770	CC.	-
_																																							
71	77	-	eT/	į	721	-	TA	TE	STAC	~	TÄAT	ere	TC1	. 111	CCT.	ATA	WE	TA	1441	_	_	_	-	_	_	÷	_	_	-	_	-	TÉT	CC A	ETA	άi	ETA	سه	201	
11	TT	17/	KT/	ÀE	TCI	44	ITA	Te	STAC	α	TÄAT	cce	ZCT	C TT	CCT.	ATA	ME.	TA	1441	_	ETTI	_	-	_	_	eci	_	_	-	_		TET	CEA	ETA	CAT	ETA	سه	CET	
T1	11	17/	KET/	ÅE1	TCI	EA.	TA)	TE	STAC	e e	TÄÄT	cce	ZCT	E TT(E T.	ATA.	MC	TAI	,	_	_	_	-	_	_		_	_	-	_	P	TET	CE A	ETA O	<u>eat</u> L	ETA M	elii L	# #	
T¶ #		M/	KET/	ÅÆT	TCI	L	ITAI	TE	STAC	G	TÄAT	e e	EC7	ETT(E C T.	ATA.	L L	N N	, ,	_	_	_	-	_	_	P	_	_	-	_	P	1161	CE A	0	eat L	eta M	r	P P	
TQ H	TŤ	1	KET/	ÅÆT	TCI R	L	ITA	76	ETAC	e G	TÄÄT	6	*	ETT	E E T.	ATA.	AAÉ1	H TA	,	_	_	_	-	_	_	eci	TCAC	_	vect	_	P		CEA 8	0	<u>eat</u> L	eta M	L	;ce1	
# #		•	•	•	R	L	•	7		G		6	•	C	6	٧	L		•	H.	e 11	#£11	P.	4	1 (C)	•	TCAC	E1/	WCT S 0	, ,	P		•	o Pai	L	M	L	,	
T1		•	•	•	R	L	•	7		G		6	•	C	6	٧	L		•	H.	_	#£11	P.	4	1 (C)	•	TCAC	E1/	WCT S 0	, ,	P		•	o Pai	L	M	L	,	2140
_		A	ATI	1	e BT	L.	i BBI	7	H CAA1	G	184	6 TAT	& BAT	E AAA	700	v Coși	L	105	,	ACC H	e 11	AC11	TCA	CC 11	H I	704	E V	ETI	VACT	AAA F	P		1111	O Boll BAT	L EAA	M M	L ATT	, aec	2 140 0
_		A	ATI	1	e BT	L.	i BBI	7	H EAA1	G TA	L AST	E TAT	BAT.	E AAA	6 700	COSI	L CET/	H rec	1110	H TATO	G I	EAA1	P TCA	AAA	H I	TEL	E N	ETI COST	VACT V	AAA F	P	R SETE	TTT	BAT BAT	EAA OTT	H BCA EQT	ATT(ACC TTCO	216 0
_		A	ATI	1	e BT	L.	i BBI	7	H EAA1	G TA	L AST	E TAT	BAT.	E AAA	6 700	COSI	L CET/	H rec	1110	H TATO	G I	EAA1	P TCA	AAA	H I	TEL	E N	ETI COST	VACT V	AAA F	P	R SETE	TTT	BAT BAT	EAA OTT	H BCA EQT	ATT(ACC TTCO	2 HEO
_		A	ATI	1	e BT	L.	i BBI	7	H EAA1	G TA	L AST	E TAT	BAT.	E AAA	6 700	COSI	L CET/	H rec	1110	H TATO	G I	EAA1	P TCA	AAA	H I	TEL	E N	ETI COST	VACT V	AAA F	P	R SETE	TTT	BAT BAT	EAA OTT	H BCA EQT	ATT(ACC TTCO	216 0
	e e	TI I	TAI	I I AM	R EAJ	L rec	e ccc	TT:	EAA1	G AT	EAST STCA	E TAT ATA	BAT.	E TTT	rec ace	V COSI SCCI	L SEAT	TCC:	F TTTE	H H TATO	G I	ACTI E CAAT STTA	TCA AGT	AAAA	H (P ACT	E N	ETI COST	VACT	AAA TTA	P DETI	R ETE	1777 444	BAT BAT	EAA OTT	M ACA TOT	ATTO TAME	ACC STEG	2160
	CE CE	M 177	TAI		EA/	L MEE	6	TT H	EATI	G ITA IAT V	EAST STCA	TAT ATA	BAT. ETA	E AAA	TOC TAC	R COSI	L DEAT	II	, 1116 1446 7	H D LATE	TAGE	EAAT ETTA	AAA P	AAAA	H I	P ACT	E V	ETI GOST TAST	MET IV	TTA	P CETT	R ICTE	111	O DOI O	EAA BTT	EGT A	ATTE	ACC STEG	2 140 2200
	CE CE	M 177	TAI		EA/	L MEE	6	TT H	EATI	G ITA IAT V	EAST STCA	TAT ATA	BAT. ETA	E AAA	TOC TAC	R COSI	L DEAT	II	, 1116 1446 7	H D LATE	TAGE	EAAT ETTA	AAA P	AAAA	H I	P ACT	E V	ETI GOST TAST	MET IV	TTA	P CETT	R ICTE	111	O DOI O	EAA BTT	EGT A	ATTE	ACC STEG	2160
CA AT TA	CC.	A I	TAI	TTO	EA/	L TEC		TT H	EAAT	G ITA	STEA STEA	TAT ATA 1	BAT. ETA B EAC	E EAA'	TOC C TAC	COGI R ATA	EST/ SEAT	H COS	TAMES	H H IATE	TAGE	ECTION OF THE COLUMN CO	AAA P	AAGO	M (ATE	P TEL	E N E N E N E N E N E N E N E N E N E N	ETC	ACT V	AAA TTA	P CETC	R ETC E	1111 AAA P	BATTA	CAA GTT 0	BCA EQT A	ATTO	ACC STEE	2160
CA AT TA	CC.	A I	TAI	TTO	EA/	L TEC		TT H	EAAT	G ITA	STEA STEA	TAT ATA 1	BAT. ETA B EAC	E EAA'	TOC C TAC	COGI B ATA	EST/ SEAT	H COS	TAMES	H H IATE	TAGE	ECTION OF THE COLUMN CO	AAA P	AAGO	M (ATE	P TEL	E N E N E N E N E N E N E N E N E N E N	ETC	ACT V	AAA TTA	P CETC	R ETC E	1111 AAA P	BATTA	CAA GTT 0	BCA EQT A	ATTO	ACC STEE	2160
	CC.	M 177	TAI	TTTT	EAL V	L TECTAL P	6	TT H	EATI EATI	G TAT	S TTET	TATA ATA T CAC	BAT. ETA B EAC	E CAA'	E TOC C TAC	COGG	L EET/ EEAT AGT/	H ABB	F AAGE	HATCH TAME	TAGE	CAAT BTTA BTTA CETA	TCA AGT S CSG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	M (P TEL	CTATE A THE COLUMN TO THE COLU	AGT	MACT I V	TTA M	P REAL CREATE	R TETE E SETT TEAM	1111 AAA P	BATE	EAA BTT •	EGT A AAT	ATTO	TAGE TITEO	2160
10 s 41 th 60	CCI G		TAI	TTO E	EAN V	L TEC	TEST COLUMN TO THE TEST COLUMN T	TT H	EATI EATI BEATI	G ITAL T ITE	CAST STCA STCT	TAT ATA 1 CTC E	BAT. ETA B EAC ETG	E CAA' ETT!	TACE	TATA	EET/ SEAT STCM AGT	H COL	F MAGE	HATTE TAKE	TAGE	ECAAT ETTA CETA	AAAA P TCA AGT S COCC R TCA	AAAA TTCC II ATCA TAS	EATER	P TEL	E IN ENTER SERVICE SER	ETC I	VACTOR OF TACK	AAA F TTA TTTA TTT	P CETT	R ETC E	**************************************	BATA CAT. Y	CAA BTT •	EGT A AAT TTA	ATTRI TAME 1 TTTRI	TABLE TITLES	2160 2260
10 s 41 th 60	CCI G		TAI	TTO E	EAN V	L TEC	TEST COLUMN TO THE TEST COLUMN T	TT H	EATI EATI BEATI	G ITAL T ITE	CAST STCA STCT	TAT ATA 1 CTC E	BAT. ETA B EAC ETG	E CAA' ETT!	TACE	TATA	EET/ SEAT STCM AGT	H COL	F MAGE	HATTE TAKE	TAGE	ECAAT ETTA CETA	AAAA P TCA AGT S COCC R TCA	AAAA TTCC II ATCA TAS	EATER	P TEL	E IN ENTER SERVICE SER	ETC I	VACTOR OF TACK	AAA F TTA TTTA TTT	P CETT	R ETC E	**************************************	BATA CAT. Y	CAA BTT •	EGT A AAT TTA	ATTRI TAME 1 TTTRI	TABLE TITLES	2160 2260 2400
10 s 41 th 60	CCI G		TAI	TTO E	EAN V	L TEC	TEST COLUMN TO THE TEST COLUMN T	TT H	EATI EATI BEATI	G ITAL T ITE	CAST STCA STCT	TAT ATA 1 CTC E	BAT. GTA B CAC RTS	E CAA' ETT/	TAC	TATA	EET.	H TACI	F MAGE	MATE O DATE O	TAGE TAGE TAGE TAGE TAGE TAGE TAGE TAGE	ECANT BTTA BTTA CETA	TCA AGT S CBG FCC R	AAAA TTCC II ATCA TAS	EATER	P TEL	E IN ENTER SERVICE SER	ETC I	VACTOR OF TACK	AAA F TTA TTTA TTT	P CETT	R ETC E	**************************************	BATA CAT. Y	CAA BTT •	EGT A AAT TTA	ATTRI TAME 1 TTTRI	TABLE TITLES	2140 2340
10 s 41 th 60	CCI G		TAI	TTO E	EAN V	L TEC	TEST COLUMN TO THE TEST COLUMN T	TT H	EATI EATI BEATI	G ITAL T ITE	CAST STCA STCT	TAT ATA 1 CTC E	BAT. GTA B CAC RTS	E CAA' ETT/	TAC	TATA	EET.	H TACI	F MAGE	MATE O DATE O	TAGE	ECANT BTTA BTTA CETA	TCA AGT S CBG FCC R	AAAA TTCC II ATCA TAS	EATER	P TEL	E IN ENTER SERVICE SER	ETC I	VACTOR OF TACK	AAA F TTA TTTA TTT	P CETT	R ETC E	**************************************	BATA CAT. Y	CAA BTT •	EGT A AAT TTA	ATTRI TAME 1 TTTRI	TABLE TITLES	2140 2240
CA S ATTA H SECT V	CCI GALL ACT	A TI	TAI	TTO	R EAA V ECG A VAI	L PECTAL PROPERTY OF THE PERTY	G G G G G G G G G G G G G G G G G G G	TT H	EAA1 ETTI BEATI GTAI	G TAC TEA	TOTAL S	TATATATA	BAT. CTA B CAC ETG H TTA	E CAA' TTC	TACE TACE TACE TACE TACE TACE TACE TACE	TATATATE	L CET/	H ASSTRACE R ATEL TO THE TAKEN	THE RESERVE IN THE RE	H D LATE	TAGE TAGE TAGE TAGE TAGE TAGE TAGE TAGE	CCT/	AAA P TCA AGT S CSG GCC R TCA AGT	AAAB TTCC E ATC: TAG: I BATI	EATER STAGE	P TEL	GAT/	ETI OCCUPANTO	TACE	AAA F	P REAL CONTROL OF THE	R SANGER OF THE	TTT AAA	BATTA B BTALL TEAT	CAA BTT	EGT A AAT TYA BAT ETA	ATTO	A CATT THAN H LAGE TEES E	2140 2340
CA S ATTA H SECT V	CCC G CTT		TEN CONTRACT	TTO E	R SEA V SECOND A VAI	L TEC	6 1881 8 1881 1731	T H T T AA	EAAT ETTI M EATI STAI GCGI	G TAIL TO THE TO TAIL	STEET AMEA	TATATATA	BAT. ETA. BEAC. ETG. HITTA	E CAA'S ETT!	TACE TACE TACE TACE TACE TACE TACE TACE	TATAL COGNI	EST/SEAT SAASTTE	H TACK	F MAGE	HATE TAKE	TAGE ATTE	ECALL ETT	AAA P TCA AGT S TCA R TCA AGT B GCA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	H (CATET TTAG IN (INTER TAGE TAGE	P TEL	E N E N E N E N E N E N E N E N E N E N	CTC CTC	THE STATE OF THE S	AAA F	P CETTO	R STITE CONTRACTOR OF THE PARTY	TTT AAA , CTT AAA	BATTA BATTA AATT	CAA BTT	BAT CTA	ATTO	E CETTE	2140 2240 2400
CA S ATTA H SECT V	CCC G CTT		TEN CONTRACT	TTO E	R SEA V SECON A VAI	L TEC	6 1881 8 1881 1731	T H T T AA	EAAT ETTI M EATI STAI GCGI	G TAIL TO THE TO TAIL	STEET AMEA	TATATATA	BAT. ETA. BEAC. ETG. HITTA	E CAA'S ETT!	TACE TACE TACE TACE TACE TACE TACE TACE	TATAL COGNI	EST/SEAT SAASTTE	H TACK	F MAGE	HATE TAKE	TAGE TAGE TAGE TAGE TAGE TAGE TAGE TAGE	ECALL ETT	AAA P TCA AGT S TCA R TCA AGT B GCA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	H (CATET TTAG IN (INTER TAGE TAGE	P TEL	E N E N E N E N E N E N E N E N E N E N	CTC CTC	THE STATE OF THE S	AAA F	P CETTO	R STITE CONTRACTOR OF THE PARTY	TTT AAA , CTT AAA	BATTA BATTA AATT	CAA BTT	BAT CTA	ATTO	E CETTE	2100 2200 2100
CA S ATTA H SECT V	CCC G CTT		TAN ITEL	TTI AM	EAN VECTO	L TEC	TESTI B B B B B B B B B B B B B B B B B B B	TO AA	EATI BATI BATI GTAI GTAI GTAI GTAI GTAI	TALL THE TAL	TAGE ACCO	TATATATA	BAT. ETA B ECAC ETA III TTA TCGA	E CAATE TOO AACT TOO E	E TREE	COGGI ECCI R ATAT 1 1 ECAI ECTI ECTI ATCI	EST.	H CONTROL TO TACE	F AAGI	HATCH D LATE OF LATE O	TAGE L G LAAAA TYTE L G LTETT	CCT/	AAA P TCA AGT S TCA GCC R TCA AGT B GCA CET	CCTIC	H (CATET STAGE SATE SATE SATE CAGA CAGA	P TRACT	E N E N E N E N E N E N E N E N E N E N	CTC CTC	THE STATE OF THE S	AAA F	P CETTO	R STITE CONTRACTOR OF THE PARTY	TTT AAA , CTT AAA	BATTA BATTA AATT	CAA BTT	BAT CTA	ATTI	E CETTE	2100 2200 2400
CA S ATTA H SECT V	CCC G CTT		TAN ITEL	TTI AM	EAN VECTO	L TEC	TESTI B B B B B B B B B B B B B B B B B B B	TO AA	EATI BATI BATI GTAI GTAI GTAI GTAI GTAI	TALL THE TAL	TAGE ACCO	TATATATA	BAT. ETA B ECAC ETA III TTA TCGA	E CAATE TOO BE T	E TREE	COGGI ECCI R ATAT 1 1 ECAI ECTI ECTI ATCI	EST.	H CONTROL TO TACE	F AAGI	HATCH D LATE OF LATE O	TAGE ATTE	CCT/	AAA P TCA AGT S TCA GCC R TCA AGT B GCA CET	CCTIC	H (CATET STAGE SATE SATE SATE CAGA CAGA	P TRACT	E N E N E N E N E N E N E N E N E N E N	CTC CTC	THE STATE OF THE S	AAA F	P CETTO	R STITE CONTRACTOR OF THE PARTY	TTT AAA , CTT AAA	BATTA BATTA AATT	CAA BTT	BAT CTA	ATTI	E CETTE	2160 2280 2400
CA 3 ATTA H 501CC V 55TT H	CCI		TAN TEN	TTO BE STORY	E STI	L TEC	G G G G G G G G G G G G G G G G G G G	TTO AA	EAAI ETT/ B EATI GTAI GTAI GCGG B EGAI D	TAIL TO SEAL STATE OF THE SEAL	TOTET AMENA	TATATATATATATATATATATATATATATATATATATA	BAT. CTA B KAC RTB III TTA L CGAT	E AAA TTT	TOCH CONTROL OF THE C	TATA	S SAAGTE K	H TACI	F MAGE	D LATE THE TARE	TAGE ATEL GAAM TYTE E I TGTM AGM CATE	EAAT BTTA	AAA P TCA AGT S TCA AGT B CCA AGT AGT AGT AGT AGT AGT AGT AGT AGT AG	CCTIC	H (CATET STAGE SATE SATE SATE CAGA CAGA	P TRACT	E N E N E N E N E N E N E N E N E N E N	CTC CTC	THE STATE OF THE S	AAA F	P CETTO	R STITE CONTRACTOR OF THE PARTY	TTT AAA , CTT AAA	BATTA BATTA AATT	CAA BTT	BAT CTA	ATTI	E CETTE	2100 2300 2400
CA 3 ATTA H 501CC V 55TT H	CCI		TAN TEN	TTO BE STORY	E STI	L TEC	BEAT STATE OF THE	T TAN TAN TAN	EAAI ETTI BETTI GTAI GTAI GETI D	G TAINT TO SEAL STATE OF THE S	TOTAL A COMMENT OF THE COMMENT OF TH	TATA ATA 1 SAB CTC E TEC ACC C CCT CGA	BATA ETA B EAC ETS H TTA TTA TCGA ECT	E CAAT ETT!	TAC	TATA	EST/SEAT	H AGE TACK	F MAGE	H DATE TAKE	TAGE L G LAAAA TYTE L G LTETT	CETAL BEAT CETAL BEAT CETAL BEAT TEST	AAA P TCA AGT S CCC R TCA AGT AGT A CCC TCA AGT AGT A CCC TCA AGT A CCC	CCTIC	H (CATET STAGE SATE SATE SATE CAGA CAGA	P TRACT	E N E N E N E N E N E N E N E N E N E N	CTC CTC	THE STATE OF THE S	AAA F	P CETTO	R STITE CONTRACTOR OF THE PARTY	TTT AAA , CTT AAA	BATTA BATTA AATT	CAA BTT	BAT CTA	ATTI	E CETTE	2140 2340 2400

Fig.5.

```
125+94. seq
116. seq
125+94. seq
116. seg
125+94. seq
116. seq
                                                           125-94. seq
                                                           TEATCATTEG TETEGGACTE CECCTTTTCAACTATEG AGCTEGGAGGT CTAAC TITCTTTCATCATCATTEGATGTGGGACTCCCCCCTTTTCAACTATGGGAGCTGGGAGGTTCTAAGGTTTCTTCTTTCA
116. seq
                                                       1350 1360 1370 1380 1380 1400 1410

340 350 7360 370 7380 390 7400

AATECAAGATGGTGGTTGGAAGAGTACAGGTTTGATGGTTTTAGATTTGATGGGGGTGACTTCCATGATGT

AATECAAG TGGTGGTTGGA GAGTACAA GTTTGATGG TY AGATTTGA GGGGTGACTTC ATGATGT

AATECAAGGTGGTGGTTGGATGAGTACAAGTTTGATGGTTCAAGATTTGAGGGGTGACTTCATGATGT

1420 1430 1440 1450 1480 1470 1480

410 420 7430 7430 7450 1480

ACACTCCCCATGGGTTGCAGGTAGCTTTTACTGGCAACTACAATGAGTACTTTGGATATGCAACTGATGT

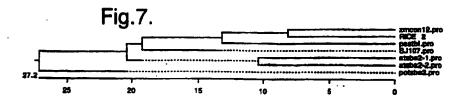
ACAC C CATGG TTGCAGGTAG TTTTACTGGCAACTACAATGA TACTTTGGATATGCAACTGATGT
125+94. seq
116. seq
125-94. seq
                                                                  CACCCATCATGGATTGCAGGTAGATTTTACEGGCAACTACAATGAATACTTTGGATATGCAACTGATGT
116. seq
                                                          1490
                                                         7480 490 7500 7510 71320 71330 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340 71340
125+94. seq
116. seq
                                                        #550 #560 #570 #580 #610 #610 GGTGAAGATGTTAGCGGAAAGCCAACATTTTGCATTCCAGTGGAAGATGTTGGATTTGATTTGATTTGATTAGCG
125+94. seq
                                                         GGTGAAGATGTTAG GGAA SCCAALA TTIGCATTCE GT GAAGATGGTGGTGTTGG TTTGATTA C
GGTGAAGATGTTAGTGGAATECCAACAGTTTGCATTCCGGTTGAAGATGGTGGTGTTGGCTTTGATTATC
1630 1840 1650 1650 1650 1650 1650 1650 1650
116. seq
                                                        GTCTCCACATEGCCATTGCCGATAAATGGATTGAGATTG
GTCTCCACATGGC TTGC GATAAATGG TTGAGATT
GTCTCCACATGGCTGATTGATAATGGGTTGAGATT
1700 1710 1720 1730
                                                                                                                                                                                                                                                           CTTAREAGEAGAGAGATEAGEACTEEAAAATGGG
TT AEAAGAGAGATEA GA TEEAAAATGGG
TATICAEAAGAGAGATGAGATTEGAAAATGGG
1740 1750 1780
125-94. seq
116. seq
                                                          125+94. seq
116. sea
                                                         CAAECTCTTGTTGGTGACAAACTATTGCATTTTGGCTGATGGACAASGACATGTACGACTTCATGGCTC
CA GC CTTGTTGGTGACAAAACTATTGCATTTTTGGCTGATGGACAAGGA ATGTA GACTTCATGGCTC
CAGGCCCTTGTTGGTGACAAAACTATTGCATTTTTGGCTGATGGACAAGGATATGTTATGACTTCATGGCTC
CAGGCCCTTGTTGGTGACAAAACTATTGCATTTTTGGCTGATGGACAAGGATATGTTATGACTTCATGGCTC
1840 1850 1860 1870
125+94. seq
116. seq
                                                       *1840 *1850 *1870 *1880 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 *1890 
125-94. seq
116. seq
                                                         125+94. seq
```

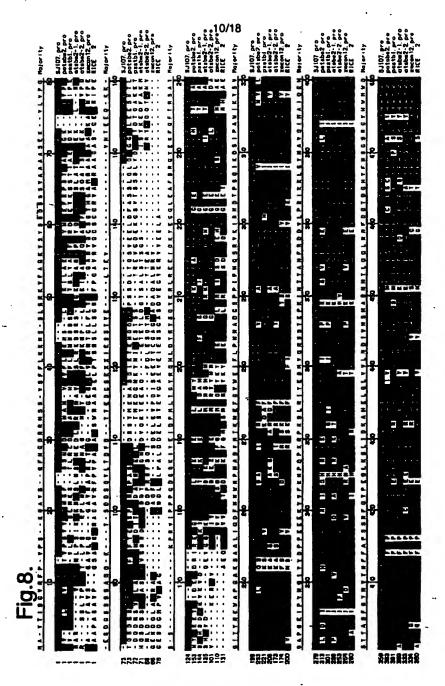
Fig.5 (Cont).

```
125+94. seq
125+94. seq
116. seq
116. seq
125-94. seq
116. seq
125-94. seq
125-94. seq
116. seq
125+94. seq
118. seq
```

Fig.6.







SUBSTITUTE SHEET (RULE 28)



SUBSTITUTE SHEET (RULE 26)

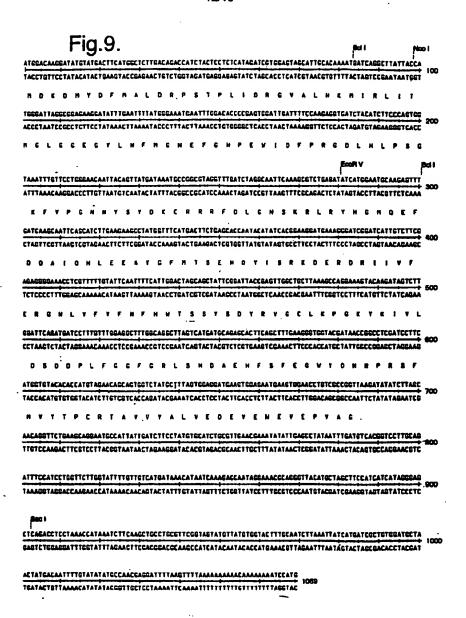


Fig. 10.

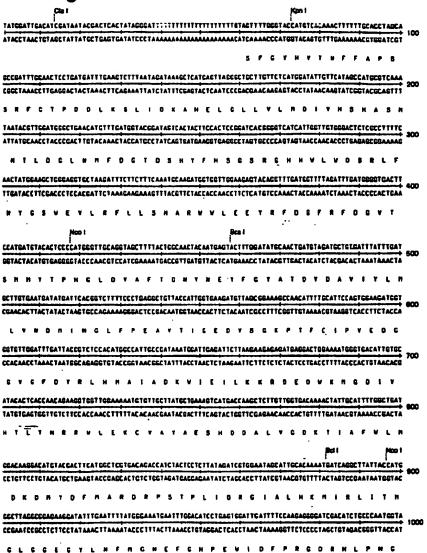
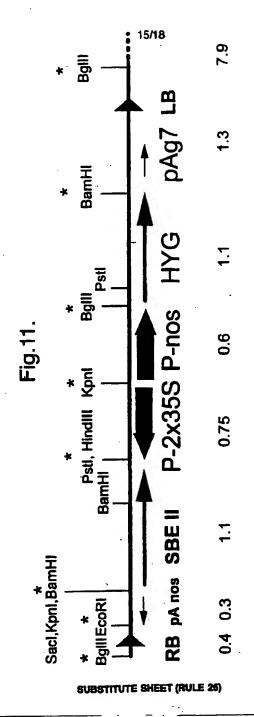


Fig. 10 (Cont).	W I
AABTAATTECABERAACAACACACACTYATGATAAATECCGTEGTAGATTTGATETAGGTGATECAGACTATCTAAGATATCATGGAATGCAABAGTTT	LA.
TTCATTAAGGTCCCTTGTTGGTGTCAATACTATTTACGGCAGCATCTAAAGTAGATCCACTACGTCTGATAGATTCTATAGTACCTTACGTTCTCAAA	+ 1100 T
K Y I P G N N H S Y D K C R R R F D L G D A D Y L R Y H G M Q E F	0
TCAGGCAATGCAACATCTTGAAGAACCCTATGCTTTCATGACTTCTGAGCACCACTATATATCAGCGAAGATGAAGGAGATCGGATCATTGTCTTTC	£
ACTCCCTTACCTTGTAGAACTTCTTCGGATACCAAACTACTGAAGACTCGTGGTCATATATAGTGCCTTCCTACTTCCTCTAGCCTAGTAACAGAAAC	+ 1200 C
Q A H Q H L E E A Y G F H T S E H O Y I 5 R K O E G O R L I V F	:
AGGGBAAACCTTGTTTTTGTATTGAACTTTCATTGEACTAACAGCTATTCAGATTACCGAGTTGGCTGCTTCAAGTCAGGAAAGTACAAQATTGTTTT	
TECCETTTEGRACAAAACATAAGTTGAAAGTAACETGATTGTCGATAAGTCTAATGGCTCAACCGACGAAGTTCAGTCCTTTCATGTTCTAACAAAA	± 1300
RGNL V F V F N F N V T N S T S D Y R V G C F K S C K Y K I V L	
ACTOBBATBATGBCTTBTTTGBAGBCTTCAACAGGCTTAGTCATBATGCCGAGCACTTCACCTTTGACGGGTGBTATGATAACCGGCCTCBBTCCTTC	T → 1400
TEASCCTACTACCGAACAACCTCCGAAGTTETCCGAATCAGTACTACCGCTCGTGAAGTBGAAACTGCCCACCATACTATTGGCCGGAGCCAGGAAG	
D S D D G L F G G F N R L S H D A E H F T F D G W Y D N R P R S F	M
	_
GOTATATGCACCATCTAGBACAGCAGTGGTCCATGCTTTAGTAGAAGATGAAGAAGAAGAAGAAAGA	→ 1600 ·
CCATATACETOSTAGATECTGTCGTCACCAGGTACGAAATCATCTTCTACTTCTCTTACTTCGTCTCTTACTTCATCTTTCACTTCACTTTGGTCGGAA	2
V Y A P B R T A V V H A L V E D E E H E A E N E V E S E V K P A :	
Personal Marcon B	
	•
GECTGAGATAGATATTTAGTAAGABBATCCCCTAAAGCAGGATGGTTAACCTGTECATCTECATTGAACGACGTATATTGAGACTTGAATTGATTTG	1800
CERCTEARATAGATATTTAGTAAGARRATCCCCTAAAGCAGGATGGTTAACCTGTECATCTECATTGAACGACGTAGACGTATATTGAEACTTGAACTTGATTGACTAACTGATTGACTAACTGACTTACCAATTGGACGTAGACGTAGACTGACCGTAGACTTGACTTAACTTAACTTAACTTAACTTAACTAAC	T → 1800
	1 1800 14
CEBACTCYATCTATAAATCATTCTCCTAGGGGATTTCGTCCTTACCAATTGGACGTAGGCGTAGGCTGACTTGCTGCATATAACTCTGAACTTAACTAAACT	1800 1800
CCENCTCTATCTATANATCATTCTCCTAGGGGATTTCGTCCTTACCAATTGGACGCGTAGGCGTAGCTTGCTGCATATAACTCTGAACTTAACTAAAC	1800 1800
CCENCTCTATCTATANATCATTCTCCTAGGGGATTTCGTCCTTACCAATTGGACGCGTAGGCGTAGCTTGCTGCATATAACTCTGAACTTAACTAAAC	ia .
CCERCTCYATCTATARATCATTCTCCTAGGCCATTTCGTCCTTACCAATTGGACACGTAGACCGTAGCTAACTTGCTGCAACTTGAACTAAACTCTGAACTAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAAA	1800 IA 1800
CCERCTCYATCTAYARATCATTCTCCTAGGGGATTTCGTCCTTACCAATTGGACACGTAGGCGTAACTTGCTGCATATAACTCTGAACTTAACTAAACT G	ia .
CCERCTCYATCTATARATCATTCTCCTAGGCCATTTCGTCCTTACCAATTGGACACGTAGACCGTAGCTAACTTGCTGCAACTTGAACTAAACTCTGAACTAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAAA	ia .
CERCTCYATCTATARATCATTCTCCTAGGGGATTTCGTCCTTACCAATTGGACACGTAGACGTAGACGTAACTTGCTGCAAATTAACTTGAACTAAACTAAACTAAACTAAACTAAACTAAACTAAACTAAACTAAACTAAACTAAACTAAACTAAACTAAAACTAAAACTAAAACTAAAACTAAAACTAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAAA	ig + 1700
CERCTCTATCTATARATCATTCTCCTAGGGGATTTCGTCCTTACCAATTGGACACGTAGACGTAGACGTAACTTGCTGCAAATTAAACTCTGAACTTAACTAAACTGCCTAACTAA	ig 1790 176 1790 17 1790 17 1800
CERCTCYATCTATARATCATTCTCCTAGGGGATTTCGTCCTTACCAATTGGACACGTAGACGTAGACGTAACTTGCTGCAAATTAACTTGAACTAAACTAAACTAAACTAAACTAAACTAAACTAAACTAAACTAAACTAAACTAAACTAAACTAAACTAAAACTAAAACTAAAACTAAAACTAAAACTAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAACTAAAAAA	ig 1790 176 1790 17 1790 17 1800
CERCTCTATCTATARATCATTCTCCTAGGGGATTTCGTCCTTACCAATTGGACACGTAGACGTAGACGTAACTTGCTGCAAATTAAACTCTGAACTTAACTAAACTGCCTAACTAA	ig 1790 176 1790 17 1790 17 1800
CERCTCTATCTATARATCATTCTCCTAGGGGATTTCGTCCTTACCAATTGGACACGTAGACGTAGACGTAACTTGCTGCAAATTAAACTCTGAACTTAACTAAACTGCCTAACTAA	ig 1790 176 1790 17 1790 17 1800
CERCTCTATCTATARATCATTCTCCTAGGGGATTTCGTCCTTACCAATTGGACACGTAGACGTAGACGTAACTTGCTGCAAATTAAACTCTGAACTTAACTAAACTGCCTAACTAA	ig 1790 176 1790 17 1790 17 1800
CERCTCTATCTATAAATCATTCTCCTAGGGGATTTCGTCCTTACCAATTGGACACGTAGACGTAGACGTAACTTGCTGCAAATTAAAETCTGAACTAAACTA	ig 1790 176 1790 17 1790 17 1800
CERCTCTATCTATARATCATTCTCCTAGGGGATTTCGTCCTTACCAATTGGACACGTAGACGTAGACGTAACTTGCTGCAAATTAAACTTGAACTTAACTAAACTAAACTGGCGATTTCGTACAATTAACTTGAACTAAACTGAACTAAACTGCGCGATAATGAAACGTAGAACAAAAACTGCTGAAAATTAAATCATTAAGCACGCCATAATGAAACACATATGAAAACCTGCCAAACTTGTAAATCATTAAGCACGCCATAATGAAACCATATGAAACCTTTTGGAGGGGTTGAAAATTTAAGCAATTAAATCGTCGCGCTACTTTTAAAATCATTTAGCAACTTTTTGAAGGGGTTGAAAATTTTAGTAAATCGTCGCCTACGTACAACTTTTAAAAAAAA	ig 1790 176 1790 17 1790 17 1800
CERCTCTATCTATARATCATTCTCCTAGGGGATTTCGTCCTTACCAATTGGACACGTAGACGTAGACGTAACTTGCTGCAAATTAAACTTGAACTTAACTAAACTAAACTGGCGATTTCGTACAATTAACTTGAACTAAACTGAACTAAACTGCGCGATAATGAAACGTAGAACAAAAACTGCTGAAAATTAAATCATTAAGCACGCCATAATGAAACACATATGAAAACCTGCCAAACTTGTAAATCATTAAGCACGCCATAATGAAACCATATGAAACCTTTTGGAGGGGTTGAAAATTTAAGCAATTAAATCGTCGCGCTACTTTTAAAATCATTTAGCAACTTTTTGAAGGGGTTGAAAATTTTAGTAAATCGTCGCCTACGTACAACTTTTAAAAAAAA	ig 1790 176 1790 17 1790 17 1800
CERCTCTATCTATARATCATTCTCCTAGGGGATTTCGTCCTTACCAATTGGACACGTAGACGTAGACGTAACTTGCTGCAAATTAAACTTGAACTTAACTAAACTAAACTGGCGATTTCGTACAATTAACTTGAACTAAACTGAACTAAACTGCGCGATAATGAAACGTAGAACAAAAACTGCTGAAAATTAAATCATTAAGCACGCCATAATGAAACACATATGAAAACCTGCCAAACTTGTAAATCATTAAGCACGCCATAATGAAACCATATGAAACCTTTTGGAGGGGTTGAAAATTTAAGCAATTAAATCGTCGCGCTACTTTTAAAATCATTTAGCAACTTTTTGAAGGGGTTGAAAATTTTAGTAAATCGTCGCCTACGTACAACTTTTAAAAAAAA	ig 1790 176 1790 17 1790 17 1800



PCT/GB97/03032

WO 98/20145

Fig. 13. ANTINAL TREMATICAME TOMORECOMMUNICATION AND CONTROL TRACTICAL AND CONTROL TRACTICAL AND CONTROL TO C	•
NOD I THE PLACE OF A A POST AT A CHARACTER AT A CHARACTER AND AND A CHARACTER	•
H # Y Y T # B B 1 R F F C A P B L T E B B L T B F H A B R R T B B B L B F	
GAMAACTTETTGETGBACAAMBAGCCTTCTMALAACGACCTTTCAMBAGAATACTTAMACTBMBAGTTTAAATTBACAABAACAAAAACTCTTGCAGAAACAAMBACTACTAGTGTA)
. PRAYMETETTETTETTEAGATATEAATTAMAAACEAETREGAEAGTTTTSAAGSAATCEEAGRITETTSGAGAAAGTETTSGAGGAMAATGATAAGAATGTTGAGGAGGA	
ACTACE MANAGEMENT AT A TRANSPORT FOR THE STREET CHARACTEC	•
· · · · · · · · · · · · · · · · · · ·	
TRANSTANAANAMBETERITTECATTECATECATECATECATECATERAAAAMETENATETANACEAMBETECATTECTECACTORCAGENAGATATATBACETANACEAMBETEATETTATATECTANACETTATATATATATATATATATATATATATATATATATA)
Mad 9	
CTHREMOSTITECCHE MELICITEMETACOMY ATTEMANATIVE DESCRIPTION AND TEMANATIVE MANUSCRIPTION AND THE ANALYTIC MANUSCRIPTION AND THE CONTROL OF THE	•
TITETTHERCANTERANCAMENATANCTYATMANAATRIGICACCTORNICTACTIONICTICATTATTHERANTTTCANCAATTRINATICCTAATGENGATERCATENCTCEBAATGA	
AAABAA THERTEAST TETTEST TATTEST AT THE ATT THE TATE OF THE TATE O	
STYTOSTETCTOROAGATTYTTTTCCCAAATAACGCATGATCATCCCACCAATTOCTCCATGATCCCACCAAGAACGCATGGATCCATCCATGCATCCAAGAATTCCATCCTGCCACCAAGACCATCCAT	•
P G T Y G (F L P R E A D E S P P I P H & S R T E 1 R R O T P S E (K + S 1 P A	
TTREATEMENTETEMENT PLANEAUCTREPANATCECATACATGCCATATACTATRATCCACCAMAGEMENAMETATETETTCAMCATCETCAGECAMAGEMENAMETATETETTCAMCATCETCAGECAMAGEMENAMETATETETTCAMCATCETCAGECAMAGEMENAMETATETETTCAMCATCETCAGECAMAGEMENAMETATETTCAMCATCAGECAMAGEMENAMETATETTCAGACAGAGECAMAGEMENAMETATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	
AMERIAGITEAAAAGITEAGICEETBACCACTTYABBETATGITACBSTATATEATACTAMBTBSTTTEETCCTCTTCATACACAAGITTGTABGAGTEBSTTTTTAGTSA	
WERPRYGAPREIPYHAEYTOPPEEERYTPEHPEPERPERA.	
TABBESTYTATBASTETEATBTTBURATBASTATBBASEC AATAATTAACACATATBECAACTTTABAST BATATGETTEE TEGEATCAAAAAGETTBCC TAGAATBE TUTTCASAT	
ATECTAMATACTTAGAGTACAACCCTACTCATCATACCTGGGTTATTAATYSTUTATAGAGTTGAAATCTCTACTATACGAAGAGCGTAGTTTTTCGAACCGATGTTACGACCAAGTCTA	_
. взусантвиванствиттантвовицтвикцачилтей:	
CATHRETATICAMEMICATICCTATTATCCTMITTTTCMITACCATATTTTTTTTTT	-
STACCEATMASTTETCSTAMSEATAATACGATELAAACCCATGSTALASISTTTGAAAAACCCTGGATCSTCGGCTAAACCTGAGGACTACCTAAACTTCAGAAATTATCTATTTCGAGI	_
HA 1 2 2 H 3 Y Y A B P 6 T H Y T H P F A P S 2 P P G T P B D L E S L I B E A H	
TREST TREST TECTOR TO THE TOTAL AT A TOTAL AT A TABLE A TREST A TRAST A TREST A CONTROL AND A TOTAL AT A TRAST A CONTROL AND A TOTAL AND A	10

Fig.13 (Cont).

